Fixpoint semantics and simulation

Hesselink, Wim H.; Thijs, Albert

Published in:
Theoretical Computer Science

Publication date:
2000

ELSEVIER

Theoretical Computer Science 238 (2000) 275–311

Theoretical Computer Science

www.elsevier.com/locate/tcs

Fixpoint semantics and simulation

Wim H. Hesselink *, Albert Thijs

Department of Mathematics and Computing Science, Rijksuniversiteit Groningen, Postbox 800, 9700 AV Groningen, The Netherlands

Received August 1997; revised February 1998
Communicated by J.W. de Bakker

Abstract

A general functorial framework for recursive definitions is presented in which simulation of a definition scheme by another one implies an ordering between the values defined by these schemes in an arbitrary model. Under mild conditions on the functor involved, the converse implication also holds: a model is constructed such that, if the values defined are ordered, there is a simulation between the definition schemes. The theory is illustrated by applications to context-free grammars, recursive procedures in imperative languages, and simulation and bisimulation of processes.

Keywords: Simulation; Fixpoint semantics; Recursion; Model

1. Introduction

The ideas we present here, came up in the search for rules to prove refinement and equivalence between recursive procedures in imperative programming languages, cf. [6]. In such a language, one may consider two procedures, $p0$ and $p1$, both defined by mutual recursion, and ask whether $p0$ refines (i.e., implements) $p1$. One of the ways to prove refinement is to give a simulation relation between the recursion schemes of $p0$ and $p1$ (a method akin to Hoare's induction rule). This observation led us to the question whether refinement can always be proved by simulation. Of course, the answer is negative since refinement may be a consequence of specific properties of program constructors like assignments and sequential composition. It turned out, however, that the answer is positive if we want refinement independent of the interpretation of the program constructors and even independent of the domain of interpretation.

* Corresponding author.
E-mail address: wim@cs.rug.nl (W.H. Hesselink).

0304-3975/00/$ - see front matter
PII: S0304-3975(98)00176-5

For the proof of this result it was convenient to take a number of abstraction steps, via program schemes and polynomial functors, to almost arbitrary functors from sets to sets.

Thus, the results we present are twofold. Firstly, soundness of inferring refinement from simulation in a very abstract setting for very arbitrary models, and secondly completeness of this proof strategy if model assumptions are discarded.

As a consequence of the abstractions, the recursive procedures we began with have been replaced by arbitrary definition schemes, and the refinement relation is replaced by an arbitrary partial order (which may be equality). The theory is applicable to many classes of recursive definitions. Of course, in specific situations the soundness result may turn out to be trivial, whereas the completeness result is irrelevant for specific models. For example, in Moschovakis' theory of functional recursion [10], the theory trivializes: simulation is not much more than renaming of functions, and completeness is irrelevant since the intended models satisfy specific laws not represented in our theory. The theory is more fruitful in nondeterministic settings: we give examples with context-free languages, recursive definitions in process algebras, and recursive procedures in imperative languages.

Summarizing, the aim of this paper is to give a general framework with a sound and complete rule (viz. simulation) to prove that one recursively defined value is greater than another one, independent of the model in which these values are defined.

We now turn to a description of recursive definitions, and of the related concepts of syntax and semantics, which culminates in a neat but abstract categorical setting. A simple recursive definition defines a value in terms of itself. Strictly speaking, a recursive definition is not a definition but a fixpoint equation, and the actual definition amounts to the choice of the best solution of the equation in terms of a relevant order. A general recursive definition defines a family of values in terms of itself (mutual recursion), but its purpose usually is to define a specific value. For example, in a context-free grammar each nonterminal is provided with a set of productions, but we are especially interested in the productions of the start symbol.

For the analysis of a recursive definition it is useful to distinguish the syntactic definition scheme from the semantic interpretation. The definition scheme consists of an index set D for the family that is to be defined, a function φ that assigns to each index an expression over D, and a special index $d \in D$ to point at the value we are interested in. At the semantic side, we have a set A of values, a function α that evaluates expressions over A to values in A, and an order to choose the best fixpoint.

For example, in the case of context-free grammars, D is the set of nonterminals, function φ provides the production rules, A is the set of languages, function α represents concatenation and union of languages, the order is inclusion, and the least fixpoint is the context-free language defined. For details, see Example 3 in Section 5.2. In Example 4 of the same section, D is a set of mutually recursive procedures, φ is the declaration of the bodies of these procedures, A is the set of monotonic predicate transformers, α provides operations for composition and nondeterministic choice, the

order is implication order, and the least fixpoint gives the weakest preconditions of the procedures.

Back to the general setup. The shape of the expressions over A must be the same as the shape of the expressions over D, since, for every "valuation" that assigns values in A to elements of D, and for every expression over D, we need a corresponding expression over A. This requirement is formalized in the condition that the set of expressions over X depends functorially on X. We can then abstract from the syntax of the expressions.

The theory therefore starts with the choice of a functor F from sets to sets. A scheme is a triple (D, φ, d) with D a set, φ a function from D to $F.D$, and d an element of D. A model is a triple $(A, \alpha, \leqslant)$ with A a set, α a function from $F.A$ to A, and $\leqslant$ an order on A, which has to satisfy some compatibility with F and α. We often use d as an abbreviation of (D, φ, d) and A as an abbreviation of $(A, \alpha, \leqslant)$. The meaning $[\![d]\!]$ of a scheme d in a model A will be defined as $f.d$ where f is the least function f from D to A with $\alpha \circ F.f \circ \varphi = f$ (which is a hylomorphism in the sense of [8]).

$$\begin{array}{ccc} D & \xrightarrow{\varphi} & F.D \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle F.f} \\ A & \xleftarrow{\alpha} & F.A \end{array}$$

The equality problem is the question whether two recursive definitions, i.e., schemes d and e, define the same values $[\![d]\!] = [\![e]\!]$. Since model A has an order, we are even more interested in the ordering problem, whether $[\![d]\!] \leqslant [\![e]\!]$. The answers to these questions may depend on the model and then domain specific methods are needed to answer them. This paper is devoted to the case that equality or order can be asserted uniformly, for a large class of models.

The key concept is that of simulation, a functorial generalization of concepts introduced by Milner [9] and Park [11]. For the definition of simulation we have to extend the functor F to binary relations between sets. It turns out that almost all important functors have at least one such relational extension. In cases with nondeterminacy like grammars and transition systems, there are usually two or three relevant relational extensions, which behave differently.

Our first main result is soundness of simulation: if d and e are schemes with a simulation from d to e, then $[\![d]\!] \leqslant [\![e]\!]$ in every well-behaved model in which d has meaning $[\![d]\!]$ and e has meaning $[\![e]\!]$. Here the term "well-behaved" has a precise definition, independent of d and e.

Our second main result is completeness in the sense that existence of a simulation follows from the assumption that $[\![d]\!] \leqslant [\![e]\!]$ holds in every well-behaved model.

Actually, we give a stronger completeness result. Let a model be called universal, if every scheme d has a meaning $[\![d]\!]$ in it, and if $[\![d]\!] \leqslant [\![e]\!]$ in it is equivalent to the existence of a simulation from d to e. We show that a universal model exists if, loosely speaking, the nondeterminacy is bounded by some cardinal number. Since

cardinal numbers can be infinite and arbitrarily large, it follows that, if a universal model does not exist, that is due to the limitations of ZF set theory.

The paper is organized as follows. In each section the results are numbered consecutively. When referring to a result from another section we add the section number to it. In Section 2 we deal with notations and some basic category theory (only functors).

In Section 3, we present the first part of the theory and we conclude with the soundness of simulation. In Section 4 we show that the functor and its relational extension together are equivalent to a functor from sets to preordered sets, which setting is more convenient for the actual constructions.

Section 5 contains the constructions of many functors with relational extensions. Here we give a number of examples and counterexamples. In particular, we show how the theory applies to context-free languages and to recursive procedures in imperative programming. We also show that the simulations form a generalization of Park's concept of bisimulation of processes, and that another relational extension induces another equivalence relation on processes.

In Section 6, we construct a univeral model and prove that an inequality $[\![d]\!] \leqslant [\![e]\!]$ in this model implies the existence of a simulation from d to e. We provide an example to show that the universal model need not be the final coalgebra of [7, 12]. Section 7 contains the conclusion and some directions for future research.

This paper is based on the first part of [14]. The functorial approach is related to the one of [7, 12], but we replace the interest in initial algebras and final coalgebras by an emphasis on the interpretation of a coalgebra in an ordered algebra.

2. Notations and basic facts

We use ordinary ZF set theory, including the axiom of choice, cf. [13]. The set of functions from a set X to a set Y is denoted by $X \to Y$. Function application is denoted by a left-associative infix dot, composition of functions by the operator $\circ$. The identity function of X is denoted by 1_X. A *functor* $F : \mathbf{Set} \to \mathbf{Set}$ is a function that assigns to every set X a set $F.X$, and to every function $f \in X \to Y$ a function $F.f \in F.X \to F.Y$, such that

$$F.1_X = 1_{F.X},$$

$$F.(f \circ g) = F.f \circ F.g \text{ for all } g \in X \to Y \text{ and } f \in Y \to Z.$$

Lemma 0. *Every functor* $\mathbf{Set} \to \mathbf{Set}$ *preserves surjective functions and bijective functions.*

Proof. Let F be such a functor. Let $f \in A \to B$ be a surjective function. By the axiom of choice there is a function $g \in B \to A$ with $f \circ g = 1_B$. Then $F.f \circ F.g = 1_{F.B}$ and therefore $F.f$ is surjective. If f is bijective we also have $g \circ f = 1_A$ and hence $F.g \circ F.f = 1_{F.A}$, so that $F.f$ is bijective. □

A relation R between sets A and B is a subset $R \subseteq A \times B$. Its converse $cv.R \subseteq B \times A$ consists of the pairs (y,x) with $(x,y) \in R$. A relation R is predominantly treated as an infix operator. We then write $x \langle R \rangle y$ for $(x,y) \in R$. The composition $R \circ S \subseteq A \times C$ of relations $R \subseteq A \times B$ and $S \subseteq B \times C$ consists of the pairs (x,z) such that $x \langle R \rangle y$ and $y \langle S \rangle z$ for some $y \in B$.

We write $=_A$ to denote the identity relation of A. A relation $R \subseteq A \times A$ is called a preorder iff $(=_A) \subseteq R$ and $R \circ R \subseteq R$. It is called an order iff moreover $R \cap cv.R = (=_A)$ (so we omit the word "partial"). The identity relation $=_A$ is also called the *discrete* preorder on A.

For a relation $R \subseteq A \times B$ and functions $f \in X \to A$ and $g \in X \to B$, we also use the *lifted* relation R given by $f \langle R \rangle g \equiv (\forall x :: f.x \langle R \rangle g.x)$. If A and B are equipped with preorders $\leqslant$, function $f \in A \to B$ is said to be monotonic iff $f.x \leqslant f.y$ in B whenever $x \leqslant y$ in A. We write **Prs** to denote the category of preordered sets with monotonic functions. Therefore, a functor $F : \mathbf{Set} \to \mathbf{Prs}$ is a functor $\mathbf{Set} \to \mathbf{Set}$ that additionally provides each set $F.X$ with a preorder $\leqslant_{F.X}$ such that every function $F.f$ is monotonic.

We write $(\forall x : P : Q)$ to mean that Q holds for all x for which P holds. Similarly, $(\exists x : P : Q)$ means that Q holds for some x for which P holds. In both cases, the range predicate P may be omitted if it is identical to *true*. So we have $(\forall x : P : Q) \equiv (\forall x :: P \Rightarrow Q)$, and $(\exists x : P : Q) \equiv (\exists x :: P \wedge Q)$, and $\neg(\forall x : P : Q) \equiv (\exists x : P : \neg Q)$. Here we have assumed that the type of x is self-evident. If that is not the case, the binding occurrence of x is written $x \in X$. We use similar notations for the quantor inf for infimum (greatest lower bound) and for λ (functional abstraction).

3. The abstract theory

Given a functor $F : \mathbf{Set} \to \mathbf{Set}$, we define in Section 3.1 the concept of relational extension. This is done axiomatically. The important concepts coalgebra, algebra, interpretation, model, pre-interpretation, post-interpretation, meaning, scheme, and refinement are all defined in Section 3.2. In Section 3.3, simulations are introduced. There we also define completeness and flatness of models, and we prove the technical lemmas for soundness. The soundness theorem itself is proved in Section 3.4. As preparations for the completeness results we there also give the definitions of universal models and separating models.

3.1. Relational extension

For relations $R \subseteq A \times B$ and $T \subseteq A' \times B'$, we define the relation $R \rightrightarrows T$ between $A \to A'$ and $B \to B'$ by

$$(f,g) \in R \rightrightarrows T \equiv (\forall x, y : x \langle R \rangle y : f.x \langle T \rangle g.y).$$

Note that this says that the product function $f \times g \in A \times B \to A' \times B'$ restricts to a function $R \to T$.

Let F be a functor **Set** → **Set**. We define a *relational extension* of F to be a function G that assigns to every pair of sets A, B, and relation $R \subseteq A \times B$, a relation $G.R \subseteq F.A \times F.B$ such that

(rel0) $(=_{F.A}) \subseteq G.(=_A)$,
(rel1) $R \subseteq S \subseteq A \times B \;\Rightarrow\; G.R \subseteq G.S$,
(rel2) $R \subseteq A \times B \wedge S \subseteq B \times C \;\Rightarrow\; G.R \circ G.S = G.(R \circ S)$,
(rel3) $(f,g) \in R \rightrightarrows T \;\Rightarrow\; (F.f, F.g) \in G.R \rightrightarrows G.T$.

These particular conditions are proof-generated: they are precisely what we need in the theory below. Yet, they are canonical to some extent: in Section 4 below, we show that relational extensions correspond to "interpolating" functors from sets to preordered sets. In Section 5 we show that many important functors have useful relational extensions. In [14], the term *relator* is used instead of relational extension. Here, we abandon the term relator to avoid confusion with the term used in [2].

In the following we fix a functor F and a relational extension G of F.

If R is a preorder on a set A, then $G.R$ is a preorder on $F.A$. In fact. $G.R$ is reflexive because of $(=_{F.A}) \subseteq G.(=_A) \subseteq G.R$. It is transitive because of $G.R \circ G.R = G(R \circ R) \subseteq G.R$.

Lemma 0. *If a function $f \in A \to B$ is monotonic with respect to preorders $\leqslant_A$ on A and $\leqslant_B$ on B, then the function $F.f \in F.A \to F.B$ is monotonic with respect to the preorders $G.(\leqslant_A)$ on $F.A$ and $G.(\leqslant_B)$ on $F.B$.*

Proof. Since monotonicity of f is the same as $(f,f) \in (\leqslant_A) \rightrightarrows (\leqslant_B)$, this follows from (rel3). □

A relation $R \subseteq A \times B$ can also be treated as a set with two projection functions $\pi_0 \in R \to A$ and $\pi_1 \in R \to B$. Then we have (obviously):

$$x \langle R \rangle y \equiv (\exists z \in R :: x \langle =_A \rangle \pi_0.z \quad \wedge \quad \pi_1.z \langle =_B \rangle y).$$

The next lemma describes a similar relationship between the set $F.R$ and the relation $G.R$.

Lemma 1. *Let $R \subseteq A \times B$, $x \in F.A$ and $y \in F.B$. Then*

$$x \langle G.R \rangle y \;\equiv\; (\exists z \in F.R :: x \langle G.(=_A) \rangle F.\pi_0.z \quad \wedge \quad F.\pi_1.z \langle G.(=_B) \rangle y).$$

Proof. ($\Rightarrow$) We define the relations $S \subseteq A \times R$ and $T \subseteq R \times B$ to consist of all pairs $(a,(a,b))$ and $((a,b),b)$, respectively, where (a,b) ranges over R. We then have $S \circ T = R$. Therefore, (rel2) implies $G.S \circ G.T = G.R$.

Now let $x \langle G.R \rangle y$. Then there exists $z \in F.R$ such that $x \langle G.S \rangle z$ and $z \langle G.T \rangle y$. We observe that $(1_A, \pi_0) \in S \rightrightarrows (=_A)$. By (rel3), it follows that $(1_{F.A}, F.\pi_0) \in G.S \rightrightarrows G.(=_A)$.

Since $x\langle G.S\rangle z$, it follows that $x\langle G.(=_A)\rangle F.\pi_0.z$. Similarly, $(\pi_1, 1_B) \in T \rightrightarrows (=_B)$, and hence $(F.\pi_1, 1_{F.B}) \in G.T \rightrightarrows G.(=_B)$, and hence $F.\pi_1.z \langle G.(=_B)\rangle y$.

($\Leftarrow$). Let $z \in F.R$ be a witness for the righthand side. It is clear that $(\pi_0, \pi_1) \in (=_R) \rightrightarrows R$. Therefore, $F.\pi_0.z\langle G.R\rangle F.\pi_1.z$ by (rel0) and (rel3). Since $(=_A) \circ R \circ (=_B) = R$, the righthand side implies $x \langle G.R\rangle y$ by (rel2). □

Remark. Let the relational extension G be called *symmetric* iff $G.(=_A)$ is symmetric for all sets A. Using Lemma 1, one can deduce from this that $G.(cv.R) = cv.(G.R)$ for all relations R. We leave this result as an aside, since the theory seems to be more fruitful in the cases where G is not symmetric. □

3.2. Coalgeberas, algebras, and interpretations

We define an *F-coalgebra* to be a pair (D, φ) where $\varphi \in D \to F.D$. We define an *F-algebra* to be a pair (A, α) where $\alpha \in F.A \to A$. See [12] for both definitions.

We define an *interpretation* of a coalgebra (D, φ) in an algebra (A, α) to be a function $f \in D \to A$ such that $\alpha \circ F.f \circ \varphi = f$. The reason for this definition is that we regard φ as a recursive definition of the elements d of D by expressions $\varphi.d$. The requirement then expresses that each value $f.d$ satisfies its recursive definition when the expressions over A are evaluated by means of α.

So, an interpretation is a fixpoint of the function P_φ from $D \to A$ to itself that is defined by $P_\varphi.f = \alpha \circ F.f \circ \varphi$. We omit the parameter α in the notation of P_φ since we often consider interpretations of different coalgebras in the same algebra (A, α). In order to guide our choice of the most adequate fixpoint we introduce a (pre-)order on the algebra A, in the following way.

A *G-premodel* is an algebra (A, α) with a preorder $\leqslant$ on A, such that the function $\alpha \in F.A \to A$ is monotonic with respct to the preorders $G.(\leqslant)$ and $\leqslant$ on $F.A$ and A, respectively. The premodel is called a *model* iff, moreover, the preorder $\leqslant$ is an order on A. Note that $G.(\leqslant)$ need not be an order on $F.A$.

For an F-coalgebra (D, φ) and a G-premodel A, we define a function $f \in D \to A$ to be a *pre-interpretation* iff f is a pre-fixpoint of P_φ, i.e., iff $P_\varphi.f \leqslant f$ for the lifted preorder on $D \to A$. Similarly, f is called a *post-interpretation* iff f is a post-fixpoint, i.e., $f \leqslant P_\varphi.f$. Since there is a good fixpoint theory for monotonic functions, we observe

Lemma 2. *Let (D, φ) be an F-coalgebra and let $(A, \alpha, \leqslant)$ be a G-premodel. Then P_φ from $D \to A$ to itself is monotonic.*

Proof. It suffices to verify that, for functions $f, g \in D \to A$:

$$
\begin{aligned}
& P_\varphi.f \leqslant P_\varphi.g \\
\equiv\ & \{\text{definition of } P_\varphi\} \\
& \alpha \circ F.f \circ \varphi \leqslant \alpha \circ F.g \circ \varphi
\end{aligned}
$$

$\Leftarrow$ {α is monotonic from $(F.A, G.(\leqslant))$ to $(A, \leqslant)$}

$F.f\langle G.(\leqslant)\rangle\ F.g$

$\Leftarrow$ {see below}

$f \leqslant g$.

Here, the last step follows from (rel3) with $R = (=_D)$ and $T = (\leqslant)$, together with (rel0) to prove $(=_{F.D}) \subseteq G.R$. □

Remark. Here we prove $f \leqslant g \Rightarrow P_\varphi.f \leqslant P_\varphi.g$ by means of a sequence of implications and equivalences with hints between braces. This linear proof format is due to Feijen, see [4, 6]. □

For an F-coalgebra (D, φ) and a G-model $(A, \alpha, \leqslant)$, we define the *meaning* $\mu\varphi$ to be the infimum (greatest lower bound) of all pre-interpretations, if that infimum exists. We say that the coalgebra (D, φ) has *meaning* in the G-model A iff $\mu\varphi$ is welldefined. Note that we do not impose completeness requirements on the G-model A.

Lemma 3. *If it is defined, the meaning $\mu\varphi$ is the least interpretation.*

Proof. Since every interpretation is a pre-interpretation and $\mu\varphi$ is the infimum of the pre-interpretations, it suffices to prove that $\mu\varphi$ is an interpretation:

$\mu\varphi = P_\varphi.\mu\varphi$

$\equiv$ {A is a G-model}

$\mu\varphi \leqslant P_\varphi.\mu\varphi \quad \wedge \quad P_\varphi.\mu\varphi \leqslant \mu\varphi$

$\Leftarrow$ {definition of $\mu\varphi$}

$P_\varphi.(P_\varphi.\mu\varphi) \leqslant P_\varphi.\mu\varphi \quad \wedge \quad P_\varphi.\mu\varphi \leqslant \mu\varphi$

$\equiv$ {P_φ is monotonic}

$P_\varphi.\mu\varphi \leqslant \mu\varphi$

$\equiv$ {definition of $\mu\varphi$}

$(\forall f : P_\varphi.f \leqslant f : P_\varphi.\mu\varphi \leqslant f)$

$\Leftarrow$ {transitivity of $\leqslant$}

$(\forall f : P_\varphi.f \leqslant f : P_\varphi.\mu\varphi \leqslant P_\varphi.f)$

$\Leftarrow$ {P_φ is monotonic}

$(\forall f : P_\varphi.f \leqslant f : \mu\varphi \leqslant f)$

$\equiv$ {definition of $\mu\varphi$}

true. $\square$

Remark. Here we prove $\mu\varphi = P_\varphi.\mu\varphi$, since we give a linear proof that the formula follows from true.

We define an *F-scheme* to be a triple (D, φ, d) such that (D, φ) is an F-coalgebra and $d \in D$. We speak of scheme d as an abbreviation of (D, φ, d). The *meaning* of scheme d in a G-model A is defined as $[\![d]\!] = \mu\varphi.d$ whenever the coalgebra (D, φ) has a meaning $\mu\varphi$ in A.

Let (D, φ, d) and (E, ψ, e) be F-schemes. We say that (D, φ, d) *refines* (E, ψ, e) in the G-model A iff both coalgebras have a meaning in A and are such that $\mu\varphi.d \leqslant \mu\psi.e$ in A. This is denoted by

$$A \models (D, \varphi, d) \sqsubseteq (E, \psi, e).$$

The purpose of this paper is to derive a sound and complete proof rule for refinement. This rule will be based on simulation, as introduced in the next section.

Example 0. Let n be a natural number. Consider the functor F given by $F.X = X^n$ (the set of the n-tuples), with the natural action on functions. For every relation $R \subseteq A \times B$, let $G.R$ be the lifted relation. Then G is easily seen to be a relational extension of F. In this case, a G-model is a triple $(A, \alpha, \leqslant)$ such that $\leqslant$ is an order on A and that $\alpha \in A^n \to A$ is monotonic. If we take $n = 3$, we have, for example, the systems of recursive equations

(i) $x = \alpha.(x, x, x)$;

(ii) $y = \alpha.(y, z, z) \wedge z = \alpha.(z, y, z)$.

If we represent the variables x, y, z, by the numbers $0, 1, 2$, respectively, the systems (i) and (ii) correspond to the F-coalgebras (D, φ) and (E, ψ) given by

$$D = \{0\}, \qquad \varphi.0 = (0, 0, 0),$$

$$E = \{1, 2\}, \qquad \psi.1 = (1, 2, 2) \wedge \psi.2 = (2, 1, 2).$$

The meaning of the coalgebra (D, φ) is the least solution of (i), and the meaning of (E, ψ) is the least solution of (ii). Let us assume that A is complete (every subset has an infimum). Then both systems have least solutions. Since every solution of (i) gives a solution of (ii) by $y, z := x, x$, we then have

$$A \models (E, \psi, 1) \sqsubseteq (D, \varphi, 0).$$

The converse relation also holds, but the proof of that fact will be a good illustration of a result in the next section. Indeed we shall prove that, for this functor, all systems of recursive equations define the same value. □

3.3. Simulation

A *G-simulation* between a coalgebra (D,φ) and a coalgebra (E,ψ) is defined to be a relation $R \subseteq D \times E$ such that $(\varphi,\psi) \in R \rightrightarrows G.R$, i.e., for all $x \in D$ and $y \in E$:

$$x \langle R \rangle y \Rightarrow \varphi.x \langle G.R \rangle \psi.y.$$

We first prove that the G-simulations form a category in the sense that identity relations are simulations and that any composition of simulations is a simulation. In fact, the identity relation $=_D$ is a G-simulation from (D,φ) to itself because of (rel0). If R and S are G-simulations from (D,φ) to (E,ψ) and from (E,ψ) to (H,χ), then $R \circ S$ is a G-simulation from (D,φ) to (H,χ), because of

$$\begin{aligned}
& x \langle R \circ S \rangle z \\
\Rightarrow\ & (\exists y :: x \langle R \rangle y \quad \wedge \quad y \langle S \rangle z) \\
\Rightarrow\ & (\exists y :: \varphi.x \langle G.R \rangle \psi.y \quad \wedge \quad \psi.y \langle G.S \rangle \chi.z) \\
\Rightarrow\ & \varphi.x \langle G.R \circ G.S \rangle \chi.z \\
\Rightarrow\ & \{ \text{ (rel2) } \} \\
& \varphi.x \langle G.(R \circ S) \rangle \chi.z.
\end{aligned}$$

We say that R is a *G-simulation of schemes* from (D,φ,d) to (E,ψ,e) iff R is a G-simulation of coalgebras from (D,φ) to (E,ψ) that contains the pair (d,e).

Remark. If the relational extension G is symmetric, every G-simulation R from a scheme d to a scheme e induces a G-simulation $cv.R$ from e to d.

We now prepare the ground for the main soundness result, which is Theorem 7 below. In Lemma 4, we show that simulation between schemes implies ordering between the values defined (i.e., the least fixpoints), if the order of model A is "complete enough". In Lemma 6, we show that a similar assertion holds for another class of models, which are called *flat*.

The cardinality of a set X is denoted by $\#X$. Let γ be a cardinal number. A preordered set $(A,\leqslant)$ is said to be *γ-complete* iff $(A,\leqslant)$ is ordered and every subset $U \subseteq A$ with cardinality $\#U \leqslant \gamma$ has an infimum.

Lemma 4. *Let (A,α) be a G-model such that the coalgebra (D,φ) has a meaning $\mu\varphi$ in A. Let $g \in E \to A$ be an interpretation of a coalgebra (E,ψ) in A. Assume that*

A is #E-complete. Let R be a G-simulation of schemes from (D,φ,d) to (E,ψ,e). Then $\mu\varphi.d \leqslant g.e$.

Proof. Since A is #E-complete, we can define a function $f \in D \to A$ by

$$f.x = (\inf\ y \in E : x \langle R \rangle y : g.y).$$

Then we have $f.x \leqslant g.y$ for all pairs x, y with $x \langle R \rangle y$. This implies $(f,g) \in R \rightrightarrows (\leqslant)$. Since R is a simulation of schemes, we have $d \langle R \rangle e$, and hence $f.d \leqslant g.e$. Therefore, it suffices to prove that $\mu\varphi \leqslant f$. This is proved in

$$\begin{aligned}
&\mu\varphi \leqslant f \\
\Leftarrow\ &\{\text{definition of } \mu\varphi\} \\
&P_\varphi.f \leqslant f \\
\equiv\ &\{\text{definition of } f\} \\
&(\forall x, y : x \langle R \rangle y : P_\varphi.f.x \leqslant g.y) \\
\equiv\ &\{\text{definition of } \rightrightarrows\ \} \\
&(P_\varphi.f, g) \in R \rightrightarrows (\leqslant) \\
\equiv\ &\{g \text{ is an interpretation, definition of } P_\varphi \text{ and } P_\psi\} \\
&(\alpha \circ F.f \circ \varphi, \alpha \circ F.g \circ \psi) \in R \rightrightarrows (\leqslant) \\
\Leftarrow\ &\{A \text{ is a } G\text{-model} : (\alpha,\alpha) \in G.(\leqslant) \rightrightarrows (\leqslant)\} \\
&(F.f \circ \varphi, F.g \circ \psi) \in R \rightrightarrows G.(\leqslant) \\
\Leftarrow\ &\{R \text{ is a } G\text{-simulation} : (\varphi,\psi) \in R \rightrightarrows G.R\} \\
&(F.f, F.g) \in G.R \rightrightarrows G.(\leqslant) \\
\Leftarrow\ &\{(\text{re13})\} \\
&(f, g) \in R \rightrightarrows (\leqslant) \\
\equiv\ &\{\text{see above}\} \\
&\text{true.} \qquad \square
\end{aligned}$$

Example 1. We come back to Example 0 with $F.X = X^n$ for a fixed number n, with its induced relational extension G. Let (D,φ,d) and (E,ψ,e) be arbitrary F-schemes. Let $\top$ be the trivial relation such that $x \langle \top \rangle y$ holds for all $x \in D$ and $y \in E$. It is

easy to verify that $\top$ is a simulation relation between the schemes, and that $cv.\top$ is a simulation in the other direction.

Assume that $(A, \alpha, \leqslant)$ is a complete G-model, so that the schemes have meaning $\mu\varphi.d$ and $\mu\psi.e$ in A. Since $\leqslant$ is an order, the Lemmas 3 and 4 imply that $\mu\varphi.d = \mu\psi.e$. So, in this case, completeness of the model implies that all schemes define the same value, which is the least solution of the analogue of system (i) of Example 0.

The next result is a technical lemma to prepare the proof of Lemma 6.

Lemma 5. *Let (E, ψ) be a coalgebra with an interpretation $g \in E \to A$ in a G-premodel $(A, \alpha, \leqslant)$. Let (D, φ) be a coalgebra and let $r \in D \to E$.*
(a) *If $F.r \circ \varphi \langle G.(=_E) \rangle \psi \circ r$ in $D \to F.E$, then $g \circ r$ is a pre-interpretation of D.*
(b) *If $\psi \circ r \langle G.(=_E) \rangle F.r \circ \varphi$ in $D \to F.E$, then $g \circ r$ is a post-interpretation of D.*

Proof. (a)

$$
\begin{aligned}
& P_\varphi.(g \circ r) \leqslant g \circ r \\
\equiv\ & \{g \text{ is an interpretation}\} \\
& P_\varphi.(g \circ r) \leqslant P_\psi.g \circ r \\
\equiv\ & \{\text{definition of } P\} \\
& \alpha \circ F.(g \circ r) \circ \varphi \leqslant \alpha \circ F.g \circ \psi \circ r \\
\Leftarrow\ & \{A \text{ is a } G\text{-premodel}\} \\
& F.g \circ F.r \circ \varphi \langle G.(\leqslant) \rangle F.g \circ \psi \circ r \\
\Leftarrow\ & \{\text{Lemma 0 with } g \in (E, =_E) \to (A, \leqslant) \text{ for } f \in A \to B\} \\
& F.r \circ \varphi \langle G.(=_E) \rangle \psi \circ r.
\end{aligned}
$$

The proof of (b) is obtained by interchanging the sides of the inequations. □

We define a G-premodel $(A, \alpha, \leqslant)$ to be *flat* iff $f \leqslant g$ holds for every coalgebra (D, φ), every post-interpretation $f \in D \to A$, and every pre-interpretation $g \in D \to A$.

Notice that, if A is a flat G-model and f is an interpretation of a coalgebra (D, φ) in A, then f is the only interpretation of (D, φ) in A, and (D, φ) has meaning $\mu\varphi = f$ in A.

Lemma 6. *Let (D, φ) and (E, ψ) be F-coalgebras with interpretations $f \in D \to A$ and $g \in E \to A$ in a flat G-premodel A. Let R be a G-simulation of schemes from (D, φ, d) to (E, ψ, e). Then $f.d \leqslant g.e$.*

Proof. Since R is a G-simulation, it follows from Lemma 1 that, for every pair $x \langle R \rangle y$, there exists an element $z \in F.R$ such that

$$\varphi.x \langle G.(=_D) \rangle F.\pi_0.z \quad \wedge \quad F.\pi_1.z \langle G.(=_E) \rangle \psi.y.$$

By the axiom of choice, it then follows that there is a function $\chi \in R \to F.R$ such that

$$(*) \qquad \varphi \circ \pi_0 \langle G.(=_D) \rangle F.\pi_0 \circ \chi \quad \wedge \quad F.\pi_1 \circ \chi \langle G.(=_E) \rangle \psi \circ \pi_1.$$

Now, (R, χ) is a coalgebra. It follows from the left-hand conjunct of $(*)$ and Lemma 5(b) that $f \circ \pi_0$ is a post-interpretation of (R, χ). It follows from the other conjunct of $(*)$ and Lemma 5(a) that $g \circ \pi_1$ is a pre-interpretation of (R, χ). Flatness of A therefore implies that $f \circ \pi_0 \leqslant g \circ \pi_1$. This proves that $f.x \leqslant g.y$ for all $x \langle R \rangle y$. □

3.4. Simulation and refinement

We can now formulate and prove our main soundness result, which is that simulation implies refinement.

To abstract from the specific simulation relation, we define the relation $\preceq_G$ between schemes by saying that $(D, \varphi, d) \preceq_G (E, \psi, e)$ holds iff there is a G-simulation of schemes R from (D, φ, d) to (E, ψ, e). Since the F-schemes with as morphisms the G-simulations between them form a category, relation $\preceq_G$ is a preorder (consequently, if G is symmetric, $\preceq_G$ is an equivalence relation).

Theorem 7. *Let (D, φ, d) and (E, ψ, e) be F-schemes with $(D, \varphi, d) \preceq_G (E, \psi, e)$. For every G-model $(A, \alpha, \leqslant)$ in which both coalgebras have meaning, and which is #E-complete or flat, we have*

$$A \models (D, \varphi, d) \sqsubseteq (E, \psi, e).$$

Proof. By assumption, there is a G-simulation R from (D, φ, d) to (E, ψ, e) and we have interpretations $\mu\varphi$ and $\mu\psi$ of (D, φ) and (E, ψ), respectively. By Lemmas 4 and 6, it follows that $\mu\varphi.d \leqslant \mu\psi.e$. □

Theorem 7 says that simulation implies refinement for a certain class of models. In Example 1 of Section 5.1 below, we show that some assumption on the models is necessary for the validity of this implication, and in the remainder of Section 5 we give a number of examples where the implication applies.

Theorem 7 may be regarded as saying that a certain proof rule is sound. This suggested the question whether it could be complete as well. Now completeness amounts to replacing the proof rule by an equivalence of the form

$$(D, \varphi, d) \preceq_G (E, \psi, e) \;\equiv\; (\forall A :: A \models (D, \varphi, d) \sqsubseteq (E, \psi, e)),$$

where A ranges over some class of models. We aim, however, at the strongest possible result, viz., that one model is enough.

Therefore, we define G-model A to be *universal* iff for all F-schemes:

$$(D,\varphi,d) \preceq_G (E,\psi,e) \;\equiv\; A \models (D,\varphi,d) \sqsubseteq (E,\psi,e).$$

Since $(\preceq_G)$ is reflexive, a necessary condition for a universal G-model is that every coalgebra has meaning in it.

For the converse implication, we define a G-premodel $(A,\alpha,\leqslant)$ to be *separating* iff for every pair of F-schemes (D,φ,d) and (E,ψ,e) and every pre-interpretation $f \in D \to A$ and post-interpretation $g \in E \to A$ with $f.d \leqslant g.e$, we have $(D,\varphi,d) \preceq_G (E,\psi,e)$.

It is easy to verify that a flat and separating G-model in which every coalgebra has meaning is universal. In Section 6.4 we use this observation to prove that, for some functors, universal models do exist.

4. Interpolating functors

In the theory of Section 3 we used a relational extension of a given functor from **Set** to itself. It turns out that in practice the relational extension together with the functor are always given by one functor **Set** $\to$ **Prs**, which has the so-called interpolation property. In this section we develop the relevant theory. It is a preparation for Section 5, where a number of examples of the theory are presented.

4.1. The construction of relational extensions

It follows from Lemma 3.1 that the relational extension G is uniquely determined by the functor F and the relations $G.(=_A)$ in the sets $F.A$. The relations $G.(=_A)$ are preorders and Lemma 3.0 implies that all functions $F.f \in F.A \to F.B$ are monotonic with respect to them. So we can form the functor $F^G : \mathbf{Set} \to \mathbf{Prs}$ which assigns to every set A the preordered set $(A, G.(=_A))$ and which acts on functions as F does. Then G is completely determined by F^G.

Conversely, let K be a functor **Set** $\to$ **Prs**. Then we can use the analogue of Lemma 3.1 to define a prescription K^+, that assigns to every relation $R \subseteq A \times B$ the relation $K^+.R \subseteq K.A \times K.B$ given by

$$x \langle K^+.R \rangle\, y \;\equiv\; (\exists u \in K.R :: x \leqslant K.\pi_0.u \quad\wedge\quad K.\pi_1.u \leqslant y),$$

where $\pi_0 \in R \to A$ and $\pi_1 \in R \to B$ are the canonical projections and $\leqslant$ denotes the preorders of $K.A$ and $K.B$.

Lemma 0. *The prescription K^+ satisfies condition* (rel0), (rel1), *and* (rel3) *of a relational extension.*

Proof. (rel0) For $R = (=_A)$, the projections π_0 and π_1 are equal, and they are bijections. Using Lemma 2.0, it follows that

$$x \langle K^+.(=_A) \rangle\, y \equiv x \leqslant_{K.A} y.$$

Since $\leqslant_{K.A}$ is reflexive, (rel0) holds.

(rel1) Let $R \subseteq S \subseteq A \times B$. Define $j \in R \to S$ to be the injection function. Let $\sigma_0 \in S \to A$ and $\sigma_1 \in S \to B$ be the projections. We have $\pi_i = \sigma_i \circ j$. We prove $K^+.R \subseteq K^+.S$ in

$$
\begin{array}{l}
\quad x \langle K^+.R \rangle y \\
\equiv \{\text{definition of } K^+\} \\
\quad (\exists u \in K.R :: x \leqslant K.\pi_0.u \quad \wedge \quad K.\pi_1.u \leqslant y) \\
\equiv \{\pi_i = \sigma_i \circ j\} \\
\quad (\exists u \in K.R :: x \leqslant K.\sigma_0.(K.j.u) \quad \wedge \quad K.\sigma_1.(K.j.u) \leqslant y) \\
\Rightarrow \{\text{take } v = K.j.u\} \\
\quad (\exists v \in K.S :: x \leqslant K.\sigma_0.v \quad \wedge \quad K.\sigma_1.v \leqslant y) \\
\equiv \{\text{definition of } K^+\} \\
\quad x \langle K^+.S \rangle y.
\end{array}
$$

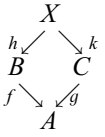
(rel3) Consider $(f,g) \in R \rightrightarrows T$. Let π'_i be the canonical projections of T. The restriction of (f,g) yields a function $j \in R \to T$ with $f \circ \pi_0 = \pi'_0 \circ j$ and $g \circ \pi_1 = \pi'_1 \circ j$. We prove $(K.f, K.g) \in K^+.R \rightrightarrows K^+.T$ by observing that, for every $x \in K.A$ and $y \in K.B$,

$$
\begin{array}{l}
\quad K.f.x \langle K^+.T \rangle K.g.y \\
\equiv \{\text{definition of } K^+\} \\
\quad (\exists v \in K.T :: K.f.x \leqslant K.\pi'_0.v \quad \wedge \quad K.\pi'_1.v \leqslant K.g.y) \\
\Leftarrow \{\text{take } v = K.j.u \text{ and use } \pi'_i \circ j = f \circ \pi_i\} \\
\quad (\exists u \in K.R :: K.f.x \leqslant K.f.(K.\pi_0.u) \quad \wedge \quad K.g.(K.\pi_1.u) \leqslant K.g.y) \\
\Leftarrow \{K.f \text{ and } K.g \text{ are monotonic}\} \\
\quad (\exists u \in K.R :: x \leqslant K.\pi_0.u \quad \wedge \quad K.\pi_1.u \leqslant y) \\
\equiv \{\text{definition of } K^+\} \\
\quad x \langle K^+.R \rangle y. \qquad \square
\end{array}
$$

In Lemma 0, condition (rel2) is still missing. In the next section we give a condition on K that implies that K^+ is a relational extension. In Section 4.3, we show that every extension can be constructed in this way.

4.2. Weak pullbacks and interpolation

Let functions $f \in B \to A$, $g \in C \to A$, $h \in X \to B$, and $k \in X \to C$ be such that $f \circ h = g \circ k$. These functions then form a so-called commuting diagram.



The diagram is called a *weak pullback diagram* iff, for every pair $b \in B, c \in C$ with $f.b = g.c$, there is an element $x \in X$ with $h.x = b$ and $k.x = c$. The diagram is called a *pullback diagram* iff, moreover, the element x is always unique. For every diagram as above, there is a canonical function from X to the subset of $B \times C$ of the pairs (b, c) with $f.b = g.c$. The diagram is a (weak) pullback iff this function is bijective (surjective).

Now let the sets A, B, C, X be preordered and let the functions f, g, h, k be monotonic. We define the diagram to be an *interpolation diagram* iff, for every pair $b \in B, c \in C$ with $f.b \leqslant g.c$ there exists an element $x \in X$ with $b \leqslant h.x$ and $k.x \leqslant c$.

Note that, if the preorders on A, B, C are discrete, the diagram is an interpolation diagram if and only if it is a weak pullback.

A functor $K : \mathbf{Set} \to \mathbf{Prs}$ is said to be *interpolating* iff it transforms every pullback diagram into an interpolation diagram.

Theorem 1. *Let $K : \mathbf{Set} \to \mathbf{Prs}$ be an interpolating functor. Then K^+ is a relational extension of K.*

Proof. By Lemma 0, it remains to verify (rel2). Let relations $R \subseteq A \times B$ and $S \subseteq B \times C$ be given. In order to relate the composition $R \circ S$ to R and S, we define the ternary relation $T \subseteq A \times B \times C$ to consist of the triples (a, b, c) with $(a, b) \in R$ and $(b, c) \in S$. Let the functions $\psi \in T \to R \circ S$ and $\psi_0 \in T \to R$ and $\psi_1 \in T \to S$ be defined by

$$\psi.(a, b, c) = (a, c) \quad \wedge \quad \psi_0.(a, b, c) = (a, b) \quad \wedge \quad \psi_1(a, b, c) = (b, c).$$

Then ψ is surjective. therefore, $K.\psi$ is surjective by Lemma 2.0.

We use π_0 and π_1 to denote the two canonical projections for each of the three binary relations $R, S, R \circ S$. The four functions $\psi_0 \in T \to R$, $\psi_1 \in T \to S$, $\pi_1 \in R \to B$, $\pi_0 \in S \to B$ form a pullback diagram: for every pair of pairs $(a, b) \in R$, $(p, q) \in S$ with $b = p$ there is a unique triple $(x, y, z) \in T$ with $(x, y) = (a, b)$ and $(y, z) = (p, q)$. Since functor K is interpolating, this implies that, for every $u \in K.R$ and $v \in K.S$,

$$K.\pi_1.u \leqslant K.\pi_0.v \quad \Rightarrow \quad (\exists t \in K.T :: u \leqslant K.\psi_0.t \quad \wedge \quad K.\psi_1.t \leqslant v).$$

By monotonicity of $K.\pi_0$ and $K.\pi_1$ and the equality $\pi_1 \circ \psi_0 = \pi_0 \circ \psi_1$, we even have the equivalence

$$(*) \qquad K.\pi_1.u \leqslant K.\pi_0.v \quad \equiv \quad (\exists t \in K.T :: u \leqslant K.\psi_0.t \quad \wedge \quad K.\psi_1.t \leqslant v).$$

The equality $K^+.R \circ K^+.S = K^+.(R \circ S)$ is proved by observing that for every pair x, y:

$$\begin{aligned}
& x \langle K^+.R \circ K^+.S \rangle y \\
\equiv\ & \{\text{composition}\} \\
& (\exists z \in K.B :: x \langle K^+.R \rangle z \quad \wedge \quad z \langle K^+.S \rangle y) \\
\equiv\ & \{\text{definition } K^+\} \\
& (\exists z \in K.B, u \in K.R, v \in K.S :: x \leqslant K.\pi_0.u \quad \wedge \quad K.\pi_1.u \leqslant z \\
& \quad \wedge \quad z \leqslant K.\pi_0.v \quad \wedge \quad K.\pi_1.v \leqslant y) \\
\equiv\ & \{\text{transitivity, good choice for } z\} \\
& (\exists u \in K.R, v \in K.S, :: x \leqslant K.\pi_0.u \quad \wedge \quad K.\pi_1.u \leqslant K.\pi_0.v \quad \wedge \quad K.\pi_1.v \leqslant y) \\
\equiv\ & \{\text{use } (*)\} \\
& (\exists u \in K.R, v \in K.S, t \in K.T :: x \leqslant K.\pi_0.u \\
& \quad \wedge \quad u \leqslant K.\psi_0.t \quad \wedge \quad K.\psi_1.t \leqslant v \quad \wedge \quad K.\pi_1.v \leqslant y) \\
\equiv\ & \{\text{calculus, choices of } u \text{ and } v\} \\
& (\exists t \in K.T :: x \leqslant K.(\pi_0 \circ \psi_0).t \quad \wedge \quad K.(\pi_1 \circ \psi_1).t \leqslant y) \\
\equiv\ & \{\pi_0 \circ \psi_0 = \pi_0 \circ \psi \text{ and } \pi_1 \circ \psi_1 = \pi_1 \circ \psi\} \\
& (\exists t \in K.T :: x \leqslant K.(\pi_0 \circ \psi).t \quad \wedge \quad K.(\pi_1 \circ \psi).t \leqslant y) \\
\equiv\ & \{\text{function } K.\psi \text{ is surjective}\} \\
& (\exists z \in K.(R \circ S) :: x \leqslant K.\pi_0.z \quad \wedge \quad K.\pi_1.z \leqslant y) \\
\equiv\ & \{\text{definition } K^+\} \\
& x \langle K^+.(R \circ S) \rangle y. \qquad \square
\end{aligned}$$

Remark. If functor K is not interpolating, the equivalence at $(*)$ can be replaced by the symbol $\Leftarrow$. So then we still have $K^+.(R \circ S) \subseteq K^+.R \circ K^+.S$.

Below we shall show that many functors **Set** → **Prs** are interpolating. Not every such functor, however, is interpolating. For example, let K be the functor that assigns to every set X the set X with the preorder $\top$ such that $x \langle \top \rangle y$ holds for all $x, y \in X$. To show that K is not interpolating, we consider a pullback diagram that consists of two disjoint nonempty subsets B and C of a set A, with the canonical injection functions f and g; we take $X = \emptyset$ with injections h and k into B and C, respectively. This is indeed a pullback diagram. It is easy to see that the K-image of the diagram is not an interpolation diagram. The functor K clearly preserves pullbacks and weak pullbacks.

There also exist interpolating functors that do not preserve weak pullbacks. For example, let S_1 and S_2 be sets with $\#S_1 = 1$ and $\#S_2 > 1$. We give both sets the preorder $\top$ introduced above. There is precisely one functor $F : \mathbf{Set} \to \mathbf{Prs}$ such that $F.\emptyset = S_2$ and $F.X = S_1$ for every nonempty X. This functor is interpolating but does not preserve weak pullbacks (it also does not preserve injective functions, compare Lemma 2.0).

4.3. Every extension comes from an interpolating functor

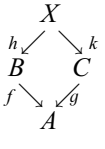
Theorem 1 shows that an interpolating functor **Set** → **Prs** gives rise to a relational extension. The converse is also true. So, we go back to the setting where G is a relational extension of an arbitray functor $F : \mathbf{Set} \to \mathbf{Set}$. As indicated above, functor F together with its extension G define a functor $F^G : \mathbf{Set} \to \mathbf{Prs}$. Below we show that F^G is interpolating. By Lemma 3.1, the relational extension $(F^G)^+$ is equal to G. Therefore, the correspondence between interpolating functors and relational extensions is bijective.

First, some convenient notation. For functions $f \in X \to A$ and $g \in X \to B$, we write $[f, g]$ to denote the relation in $A \times B$ that consists of the pairs $(f.x, g.x)$ with $x \in X$. We clearly have $(f, g) \in (=_X) \rightrightarrows [f, g]$. By (rel3) and (rel0) it follows that $(F.f, F.g) \in (=_{F.X}) \rightrightarrows G.[f, g]$. This implies

Lemma 2. $[F.f, F.g] \subseteq G.[f, g]$.

Now we can prove, as announced above:

Lemma 3. *The functor F^G is interpolating.*



Proof. It suffices to prove that F^G transforms any weak pullback diagram into an interpolation diagram. So, consider a weak pullback diagram as in Section 4.2. Let $b \in F.B$ and $c \in F.C$ be such that $F.f.b \langle G.(=_A) \rangle F.g.c$. We now have to prove the existence of $x \in F.X$ with $b \langle G.(=_B) \rangle F.h.x$ and $F.k.x \langle G.(=_C) \rangle c$. Let $R \subseteq B \times C$

consist of the pairs (p,q) with $f.p = g.q$. Since the diagram is a weak pullback diagram, there is a surjective function $m \in X \to R$ with $h = \pi_0 \circ m$ and $k = \pi_1 \circ m$. By Lemma 2.0, function $F.m$ is surjective. So, it suffices to prove the existence of $z \in F.R$ with $b \langle G.(=_B) \rangle F.\pi_0.z$ and $F.\pi_1.z \langle G.(=_C) \rangle c$. By Lemma 3.1, it therefore remains to prove $b \langle G.R \rangle c$. We now observe that

$$[1_B, f] \circ (=_A) \circ [g, 1_C] = R,$$

and hence by (rel2) and Lemma 2,

$$[1_{F.B}, F.f] \circ G.(=_A) \circ [F.g, 1_{F.C}] \subseteq G.R.$$

Therefore $F.f.b \langle G.(=_A) \rangle F.g.c$ implies $b \langle G.R \rangle c$. □

We have thus shown the converse of Theorem 1: if G is a relational extension of an arbitrary functor F from sets to sets, there is an interpolating functor K from sets to preordered sets such that $G = K^+$. Since it is easier to construct interpolating functors than relational extensions, we regard Theorem 1 as more important than its converse.

4.4. Interpolating functors as a composition

The next results will be used to compute relational extensions and to construct models for them. If K is an interpolating functor **Set** → **Prs** we speak of a model for K where a K^+-model is meant. Implicitly, an arbitrary functor $F : \mathbf{Set} \to \mathbf{Set}$ is also regarded as a functor **Set** → **Prs**, where every set $F.X$ is provided with the discrete preorder $=_{F.X}$. Then functor F is interpolating if and only if it preserves weak pullbacks.

Remark. It follows from Lemma 2.0 that a functor that preserves pullback diagrams also preserves weak pullback diagrams. Since preservation of weak pullbacks is the weaker property, it is used as the hypothesis in the Lemmas 4 and 5 below. We use these lemmas, however, only for functors that actually preserve pullbacks.

Lemma 4. *Let $H : \mathbf{Set} \to \mathbf{Set}$ be a functor that preserves weak pullback diagrams. Then H is interpolating and hence has a relational extension H^+. Let $K : \mathbf{Set} \to \mathbf{Prs}$ be an interpolating functor. Then $K \circ H$ is interpolating and $(K \circ H)^+.R = K^+.(H^+.R)$ for every relation R.*

Proof. We need only prove the equality between the relations. Consider a relation $R \subseteq A \times B$. We use $\pi_0 \in R \to A$, $\pi_1 \in R \to B$, $\tau_0 \in H^+.R \to H.A$, $\tau_1 \in H^+.R \to H.B$ to denote the canonical projections. Since $H.A$ and $H.B$ have the discrete orders, relation $H^+.R$ consists of the pairs $(x,y) \in H.A \times H.B$ such that there exists $u \in H.R$ with $x = H.\pi_0.u$ and $H.\pi_1.u = y$. So, there is a surjective function $f \in H.R \to H^+.R$ such that $\tau_i \circ f = H.\pi_i$ for $i = 0, 1$. By Lemma 2.0. it follows that $K.f$ is also surjective.

Now we conclude by observing

$$
\begin{aligned}
& x \langle K^+.(H^+.R) \rangle y \\
\equiv\ & \{\text{definition of } K^+\} \\
& (\exists u \in K.(H^+.R) : x \leqslant K.\tau_0.u \quad \wedge \quad K.\tau_1.u \leqslant y) \\
\equiv\ & \{K.f \in K.(H.R) \to K.(H^+.R) \text{ is surjective}\} \\
& (\exists v \in K.(H.R) : x \leqslant K.\tau_0.(K.f.v) \quad \wedge \quad K.\tau_1.(K.f.v) \leqslant y) \\
\equiv\ & \{\tau_i \circ f = H.\pi_i, \text{ calculus}\} \\
& (\exists v \in (K \circ H).R : x \leqslant (K \circ H).\pi_0.v \quad \wedge \quad (K \circ H).\pi_1.v \leqslant y) \\
\equiv\ & \{\text{definition of } (K \circ H)^+\} \\
& x \langle (K \circ H)^+.R \rangle y. \qquad \square
\end{aligned}
$$

Lemma 5. *Let $(A, \beta, \leqslant)$ be a model for an interpolating functor K and let $(A, \gamma, \leqslant)$ be a model for a functor $H : \mathbf{Set} \to \mathbf{Set}$ that preserves weak pullbacks. Putting $\alpha = \beta \circ K.\gamma$, we have that $(A, \alpha, \leqslant)$ is a model for $K \circ H$.*

Proof. Since $(\leqslant_A)$ is an order on A, it suffices to verify monotonicity of α. The function $\gamma \in H.A \to A$ is monotonic for the preorders $H^+.(\leqslant_A)$ and $(\leqslant_A)$. By Lemma 3.0 this implies that $K.\gamma$ is monotonic for $K^+.(H^+.(\leqslant_A))$ and $K^+.(\leqslant_A)$. Since function β is monotonic for $K^+.(\leqslant_A)$ and $(\leqslant_A)$, it follows with Lemma 4 that α is monotonic for $(K \circ H)^+.(\leqslant_A)$ and $(\leqslant_A)$, as required. $\square$

5. Concrete functors and examples

In Section 5.1 we present a number of power set functors, which form typical examples of interpolating functors. Here we also give some toy examples of the theory of Section 3. In Section 5.2 we introduce string functors. Even more important than strings are sets of strings, languages. For this purpose a string functor is combined with a power set functor. It is shown that context-free grammars and mutually recursive procedures in an imperative language are both examples of the theory.

Section 5.3 treats polynomial functors and their nondeterministic relatives. Here, we show that Park's concept of bisimulation of processes (or transition systems) is a special case of our concept of simulation, but that simulation may also induce an equivalence relation different from bisimulation. In Section 5.4 we introduce the concept of bounded spread for functors. This concept is needed for the construction of a universal model in Section 6.4.

5.1. The power functors

Let $\mathscr{P}ow : \mathbf{Set} \to \mathbf{Set}$ be the functor that assigns to every set A the set $\mathscr{P}ow.A$ of all subsets of A, and that assigns to a function $f \in A \to B$ the function f_s that assigns to a subset $U \subseteq A$ the image $im(f|U) \subseteq B$. We also define the functors $\mathscr{P}oi$, $\mathscr{P}oc : \mathbf{Set} \to \mathbf{Prs}$ that treat sets and functions in the same way as $\mathscr{P}ow$, but are such that $\mathscr{P}oi.X$ is ordered by inclusion ($\subseteq$) and that $\mathscr{P}oc.X$ is ordered by containment ($\supseteq$). It is easy to see that, in either case, all functions f_s are monotonic. Therefore, they are indeed both functors $\mathbf{Set} \to \mathbf{Prs}$.

We also introduce several related functors. We fix an infinite cardinal number κ and define the subfunctors $\mathscr{P}owk$, $\mathscr{P}own$, $\mathscr{P}ownk$ of $\mathscr{P}ow$ by

$$U \in \mathscr{P}owk.X \quad \equiv \quad \#U < \kappa,$$

$$U \in \mathscr{P}own.X \quad \equiv \quad 0 < \#U,$$

$$U \in \mathscr{P}ownk.X \quad \equiv \quad 0 < \#U < \kappa;$$

it is easy to see that indeed, for $f \in X \to Y$, the function $f_s \in \mathscr{P}ow.X \to \mathscr{P}ow.Y$ restricts correctly to $\mathscr{P}owk.X \to \mathscr{P}owk.Y$, etc.

We provide the results of these functors with the inclusion or containment ordering, by replacing the w by i or c. This gives us six other functors $\mathbf{Set} \to \mathbf{Prs}$. For example, $\mathscr{P}ocn.X$ is the set of nonempty subsets of X, ordered by $\supseteq$.

Lemma 0. *The 12 functors $\mathscr{P}ow$, $\mathscr{P}owk$, $\mathscr{P}own$, $\mathscr{P}ownk$, $\mathscr{P}oi$, $\mathscr{P}oik$, $\mathscr{P}oin$, $\mathscr{P}oink$, $\mathscr{P}oc$, $\mathscr{P}ock$, $\mathscr{P}ocn$, $\mathscr{P}ocnk$ are interpolating.*

Proof. We first treat the functor $\mathscr{P}oi$. Let $f \in B \to A$, $g \in C \to A$, $h \in X \to B$, $k \in X \to C$ form a pullback diagram. Let $U \in \mathscr{P}oi.B$ and $V \in \mathscr{P}oi.C$ be such that $f_s.U \subseteq g_s.V$. In order to show that $\mathscr{P}oi$ is interpolating, we have to exhibit a subset $W \in \mathscr{P}ow.X$ with $U \subseteq h_s.W$ and $k_s.W \subseteq V$. We try the candidate

$$W = \{x \in X | h.x \in U \quad \wedge \quad k.x \in V\}.$$

Clearly, $h_s.W \subseteq U$ and $k_s.W \subseteq V$. It remains to show that $U \subseteq h_s.W$. Let $u \in U$. Since $f_s.U \subseteq g_s.V$, there exists $v \in V$ with $f.u = g.v$. Since the diagram is a pullback, there is a (unique) $x \in X$ with $h.x = u$ and $k.x = v$. Therefore, $x \in W$ and hence $u \in h_s.W$. This proves $U \subseteq h_s.W$. So, $\mathscr{P}oi$ is interpolating. Note that we even have $h_s.W = U$.

The case of $\mathscr{P}oc$ follows by symmetry from $\mathscr{P}oi$.

In order to prove that $\mathscr{P}ow$ is interpolating, it suffices to show that, if $f_s.U = g_s.V$, then $h_s.W = U$ and $k_s.W = V$. Above we noted that $h_s.W = U$. Since the definition of W is symmetric in U and V, the additional assumption $f_s.U = g_s.V$ also implies $k_s.W = V$.

In order to prove that $\mathscr{P}owk$ is interpolating, it suffices to show that, if $\#U < \kappa$ and $\#V < \kappa$, then $\#W < \kappa$. Since the diagram is a pullback, the function $(\lambda w :: (h.w, k.w))$

is an injection of W into $U \times V$. Therefore, $\#W \leqslant \#U \times \#V$. Since κ is an infinite cardinal number, and $\#U < \kappa$ and $\#V < \kappa$, this implies $\#W < \kappa$ (by [13], 10.39).

In order to show that $\mathscr{P}own$ is interpolating, it suffices to observe that, if U is nonempty, then W is nonempty since $h_s.W = U$.

The other seven cases use easy combinations of the above arguments. □

Remark. The 12 functors do not preserve pullback diagrams: the set W in the above proof is usually not the only solution. They do preserve weak pullbacks.

Now that we have several interpolating functors, we can determine the corresponding relational extensions $\mathscr{P}ow^+$, etc., cf. Section 4.1.

Lemma 1. *Let $R \subseteq A \times B$. Let $u \subseteq A$ and $v \subseteq B$.*

(a) $u \langle \mathscr{P}oi^+.R \rangle v \equiv (\forall x \in u :: (\exists y \in v :: x \langle R \rangle y))$;

(b) $u \langle \mathscr{P}oc^+.R \rangle v \equiv (\forall y \in v :: (\exists x \in u :: x \langle R \rangle y))$;

(c) $\mathscr{P}ow^+.R = \mathscr{P}oi^+.R \cap \mathscr{P}oc^+.R$;

(d) $\mathscr{P}oik^+.R$, $\mathscr{P}oin^+.R$, $\mathscr{P}oink^+.R$ *are the restrictions of* $\mathscr{P}oi^+.R$; *the same holds if* i *is replaced by* c *or* w.

Proof. (a) By definition, $u \langle \mathscr{P}oi^+.R \rangle v$ is equivalent to

(a′) $(\exists w \in \mathscr{P}oi.R :: u \subseteq (\pi_0)_s.w \quad \wedge \quad (\pi_1)_s.w \subseteq v)$.

This implies the right-hand side of (a), since for every element $x \in u$ there is an element y with $(x, y) \in w$, and then $x \langle R \rangle y$ and $y \in v$. Conversely, the right-hand side of (a) implies the existence of a function $s \in u \to v$ with $x \langle R \rangle s.x$ for all $x \in u$. Let $w \subseteq R$ consist of all pairs $(x, s.x)$ with $x \in u$. Then w is a witness for condition (a′).

(b) Follows from (a) by symmetry.

(c) By definition, $u \langle \mathscr{P}ow^+.R \rangle v$ is equivalent to

(c′) $(\exists w \in \mathscr{P}ow.R :: u = (\pi_0)_s.w \quad \wedge \quad (\pi_1)_s.w = v)$.

Since (c′) implies (a′), we have $\mathscr{P}ow^+.R \subseteq \mathscr{P}oi^+.R$. By symmetry, we also have $\mathscr{P}ow^+.R \subseteq \mathscr{P}oc^+.R$. Therefore, $\mathscr{P}ow^+.R$ is contained in the intersection. Conversely, let (u, v) be an element of the righthand side of (c). We then have functions $s \in u \to v$ and $t \in v \to u$, such that $x \langle R \rangle s.x$ and $t.y \langle R \rangle y$ for all $x \in u$ and $y \in v$. Let w be the set of all these pairs $(x, s.x)$ and $(t.y, y)$. Then w is a witness for (c′).

(d) It is clear that $\mathscr{P}oik^+.R \subseteq \mathscr{P}oi^+.R \cap (\mathscr{P}oik.A \times \mathscr{P}oik.B)$. For the converse inclusion, let (u, v) be an element of the righthand side. We take w as constructed in the proof of part (a). Now it suffices to observe that $\#w = \#u$, so that $w \in \mathscr{P}oik.R$ since $u \in \mathscr{P}oik.A$. The other eight cases are similar. □

Example 0. We can now show that, if G is a relational extension, then $G.(=_A)$ need not be symmetric, and that for an order $\leqslant$ the relation $G.(\leqslant)$ need not be an order and $cv.(G.(\leqslant))$ may differ from $G.(\geqslant)$.

In fact, we take $G = \mathscr{P}oi^+$ and use part (a) of the above lemma. The first assertion follows from

$$u \langle \mathscr{P}oi^+.(=_A) \rangle v \quad \equiv \quad u \subseteq v.$$

For the other two assertions we take the standard order $\leqslant$ on the set of the natural numbers. Relation $\mathscr{P}oi^+.(\leqslant)$ is not an order because of

$$\{1,2,3\} \langle \mathscr{P}oi^+.(\leqslant) \rangle \{1,3\},$$

$$\{1,3\} \langle \mathscr{P}oi^+.(\leqslant) \rangle \{1,2,3\}.$$

Relation $cv.(\mathscr{P}oi^+.(\leqslant))$ differs from $\mathscr{P}oi^+.(\geqslant)$ because of

$$\{2\} \langle \mathscr{P}oi^+.(\leqslant) \rangle \{1,3\}, \qquad \neg(\{1,3\} \langle \mathscr{P}oi^+.(\geqslant) \rangle \{2\}).$$

Example 1. We show that in Theorem 3.7, some condition on the model $(A, \alpha, \leqslant)$ is necessary.

For this purpose we take the functor $F = \mathscr{P}own$ with its relational extension $G = \mathscr{P}own^+$. Let (D, φ) be the coalgebra with $D = \{\mathtt{d}, \mathtt{e}\}$ and $\varphi \in D \to F.D$ given by $\varphi.\mathtt{d} = \{\mathtt{d}, \mathtt{e}\}$ and $\varphi.\mathtt{e} = \{\mathtt{d}\}$. We choose relation $R = D \times D$. By Lemma 1, we have $G.R = F.D \times F.D$. Therefore, R is a G-simulation and we have

$$(D, \varphi, \mathtt{d}) \preceq_G (D, \varphi, \mathtt{e}).$$

We now construct a G-model A. We choose a natural number $n \geqslant 2$ and let A consist of the natural numbers $< n$ with the discrete order $(=)$. Using Lemma 1, we get that the relation G.$(=)$ on $F.A$ is the identity relation. Therefore, $(A, \alpha, =)$ is a G-model for every function $\alpha \in F.A \to A$. We choose α such that

(i) $\alpha.\{x\} \neq x$ for every $x \in A$,
(ii) $\alpha.U = 0$ whenever $\#U = 2$.

We now determine all pre-interpretations $f \in D \to A$:

$$\begin{aligned}
& f \text{ is a pre-interpretation} \\
\equiv\ & \{\text{definitions, the order on } A \text{ is discrete}\} \\
& \alpha \circ f_s \circ \varphi = f \\
\equiv\ & \{\text{definition of } \varphi, \text{ equality of functions}\} \\
& \alpha.\{f.\mathtt{d}, f.\mathtt{e}\} = f.\mathtt{d} \quad \wedge \quad \alpha.\{f.\mathtt{d}\} = f.\mathtt{e} \\
\equiv\ & \{f.\mathtt{d} = f.\mathtt{e} \text{ contradicts the second conjunct by (i)}\} \\
& f.\mathtt{d} \neq f.\mathtt{e} \quad \wedge \quad \alpha.\{f.\mathtt{d}, f.\mathtt{e}\} = f.\mathtt{d} \quad \wedge \quad \alpha.\{f.\mathtt{d}\} = f.\mathtt{e} \\
\equiv\ & \{\text{(ii)}\} \\
& f.\mathtt{d} = 0 \quad \wedge \quad f.\mathtt{e} = \alpha.\{0\} \neq 0.
\end{aligned}$$

This proves that there is precisely one pre-interpretation, which indeed is an interpretation, and which satisfies $f.\mathsf{d} \neq f.\mathsf{e}$. Then $\mu\varphi = f$ and we have

$$A \not\models (D, \varphi, \mathsf{d}) \sqsubseteq (D, \varphi, \mathsf{e}).$$

Example 2. Let $(A, \leqslant)$ be a complete lattice. Then we have functions sup and inf $\in \mathscr{P}ow.A \to A$ that assign to a subset $u \subseteq A$ the supremum $\sup.u$ and the infimum $\inf.u$. It follows easily from Lemma 1 that

$$u \langle \mathscr{P}oi^{+}.\leqslant \rangle v \;\Rightarrow\; \sup.u \leqslant \sup.v,$$

$$u \langle \mathscr{P}oc^{+}.\leqslant \rangle v \;\Rightarrow\; \inf.u \leqslant \inf.v.$$

Therefore, $(A, \sup)$ is a complete model for $\mathscr{P}oi$ and $(A, \inf)$ is a complete model for $\mathscr{P}oc$. Both are complete models for $\mathscr{P}ow$.

5.2. String Functors

Let $(-)^* : \mathbf{Set} \to \mathbf{Set}$ be the functor that assigns to every set X the set X^* of finite strings over X and that extends functions accordingly. This functor preserves pullback diagrams. In fact, let a pullback diagram be given as in Section 4.2. If $u \in B^*$ and $v \in C^*$ have the same image in A^*, then the strings u and v have the same length, say n, and for every index $i < n$ the i-th elements $u.i$ and $v.i$ have the same image in A. Since the diagram is a pullback, there exist unique elements $x.i \in X$ with images $u.i$ and $v.i$ in B and C. Therefore $(x.0, \ldots, x.(n-1))$ is the unique element of X^* with images u and v.

Example 3. The language generated by a context-free grammar. Let T be a set of terminal symbols. For any set X, let $X + T$ be the disjoint union of X and T. The functor $X \mapsto X + T$ is easily seen to preserve pullbacks. Therefore, the functor $X \mapsto (X + T)^*$ also preserves pullbacks. By Lemma 0 and Lemma 4.4, it follows that the functor F given by $F.X = \mathscr{P}oi.(X + T)^*$ is interpolating.

It is this functor F that is used in the definition of context-free grammars and languages. In fact, a context-free grammar over T is given by a set D of nonterminal symbols, a start symbol $d \in D$ and a function $\varphi \in D \to F.D$ that assigns to every nonterminal a set of productions. Thus, a context-free grammar is an F-scheme (D, φ, d). If we want to specify that every nonterminal has a nonempty finite set of productions, we replace $\mathscr{P}oi$ by $\mathscr{P}oink$ where $\kappa = \omega$.

In this case we are only interested in interpretations of F-schemes in the ordered F-algebra that consists of the languages over T, i.e., $A = (\mathscr{P}oi.T^*, \alpha, \subseteq)$ where $\alpha \in F.A \to A$ is constructed as follows. Let $\beta \in (A + T)^* \to A$ be given by concatenation of languages and insertion of T into singleton strings and languages. The function α is defined by $\alpha.U = (\bigcup u \in U :: \beta.u)$. It can be shown that A is a model and that $\mu\varphi.d$ is indeed the language generated by the grammar (D, φ, d). The model A is complete. So, Theorem 3.7 is applicable. This boils down as follows.

If R is a binary relation between sets D and E, we define relation R^* between $(D+T)^*$ and $(E+T)^*$ by

$$u \langle R^* \rangle v \quad \equiv \quad \#u = \#v \quad \wedge \quad (\forall i :: u.i = v.i \in T \quad \vee \quad u.i \langle R \rangle v.i).$$

It follows from Lemma 1(a) and Lemma 4.4 that such a relation is a simulation between grammars (D, φ, d) and (E, ψ, e) iff $d \langle R \rangle e$ holds and, for every pair of nonterminals u and v with $u \langle R \rangle v$ and every production $x \in \varphi.u$, there is a production $y \in \psi v$ with $x \langle R^* \rangle y$. If there is such a simulation the language generated by (D, φ, d) is contained in the language of (E, ψ, e). This is presumably well known, and it is easy to prove in the concrete setting.

Obviously, not every inclusion can be proved in this way. For instance, the regular grammars $d \to \varepsilon | td$ and $e \to \varepsilon | et$, where t is a terminal symbol, generate the same language t^*, but there is no simulation between the grammars.

Example 4. In [6], the semantics of imperative programs is expressed in terms of monotonic predicate transformers on a fixed state space. These predicate transformers form a complete lattice MT with respect to the implication order. Nondeterminate choice between a nonempty set of commands corresponds to the infimum (conjunction) of the corresponding predicate transformers. The semantics of the simple commands is given by the weakest precondition function $wp \in S \to MT$ where S is the set of simple commands. Mutually recursive procedures are given by a declaration

$$\mathbf{body} \in H \to \mathscr{P}ocn.(H+S)^*,$$

where H is the set of procedure names. Here, we use the functor $\mathscr{P}ocn$, since [6] only allows nonempty sets, and since the containment order allows the use of the infimum in the model, see Example 2 above. In fact, using Example 2 and Lemma 4.5 one can provide MT with the structure of a model for the functor $X \mapsto \mathscr{P}ocn.(X+S)^*$. Then the meanings $wp.h$ of procedure names h are determined by the least fixpoint $\mu\mathbf{body} \in H \to MT$.

In this case, a scheme (D, φ, d) is a declaration of mutually recursive procedures $u \in D$ with bodies $\varphi.u \in Pocn.(D+S)^*$, and a main procedure $d \in D$. Given two schemes (D, φ, d) and (E, ψ, e), a relation R between D and E is extended to a relation R^* in the same way as in Example 3 above. Here it follows from Lemmas 1(b) and 4.4 that R is a simulation between the schemes iff $d \langle R \rangle e$ holds and, for every pair of procedure names $u \in D$ and $v \in E$ with $u \langle R \rangle v$ and every string $y \in \psi.v$, there is $x \in \varphi.u$ with $x \langle R^* \rangle y$. Now, Theorem 3.7 implies that, if there is a simulation R between the schemes, procedure d is a refinement of e in the sense that $[wp.d.p \Rightarrow wp.e.p]$ for every postcondition p.

5.3. Polynomial functors

We fix a set Op of operator symbols and a function $\eta \in Op \to Card$ that assigns to every operator a cardinal number, its "arity". A cardinal number γ is a (well-ordered)

set with cardinality γ, cf. [13]. So, for a set A, we can identify A^γ with the set of functions $\gamma \to A$.

For a set A, we define the set $\mathcal{O}.A$ to be the disjoint union $(\sum p \in Op :: A^{\eta.p})$. This set consists of the pairs (p,I) with $p \in Op$ and $I \in \eta.p \to A$. For a function $f \in A \to B$, we define $\mathcal{O}.f : \mathcal{O}.A \to \mathcal{O}.B$ by $\mathcal{O}.f.(p,I) = (p, f \circ I)$. In this way, $\mathcal{O}$ is a functor **Set** $\to$ **Set**. A functor constructed in this way is called a polynomial functor.

An $\mathcal{O}$-algebra is a pair (A,α) with $\alpha \in \mathcal{O}.A \to A$. Since $\mathcal{O}.A$ is the disjoint union of sets $A^{\eta.p}$, function α is uniquely determined by its restrictions $\alpha_p \in A^{\eta.p} \to A$, the so-called operations. Therefore, if the sets Op and the arities $\eta.p$ are finite, $\mathcal{O}$-algebras are just $\sum$-algebras in the classical sense of [3].

Lemma 2. *Every polynomial functor $\mathcal{O}$ preserves pullback diagrams.*

Proof. Consider a pullback diagram as in Section 4.2. Let $(p,I) \in \mathcal{O}.B$ and $(q,J) \in \mathcal{O}.C$ have the same image in $\mathcal{O}.A$. Then $p = q$ and $f \circ I = g \circ J$. So, for every $r \in \eta.p$, we have $f.(I.r) = g.(J.r)$. Since the diagram is a pullback, every r has a unique $K.r \in X$ with $h.(K.r) = I.r$ and $k.(K.r) = J.r$. So, there is a unique pair in $\mathcal{O}.X$ with images $(p,I) \in \mathcal{O}.A$ and $(q,j) \in \mathcal{O}.B$, namely (p,K). □

We may regard functor $\mathcal{O}$ as a functor **Set** $\to$ **Prs**. Since it preserves pullbacks, this functor is interpolating. It is straightforward to verify that the associated relational extension $\mathcal{O}^+$ is characterized by

Lemma 3. *For a relation $R \subseteq A \times B$ and elements $(p,I) \in \mathcal{O}.A$ and $(q,J) \in \mathcal{O}.B$, we have*

$$(p,I) \langle \mathcal{O}^+.R \rangle (q,J) \quad \equiv \quad p = q \quad \wedge \quad I \langle R \rangle J.$$

Note that, by convention, $I \langle R \rangle J$ expresses $(I,J) \in (=)\rightrightarrows R$.

We now combine the polynomial functor $\mathcal{O}$ with the power functors to model non-determinacy. We define the functors $\mathcal{2}w$, $\mathcal{2}wk$, $\mathcal{2}i$, $\mathcal{2}ik : \mathbf{Set} \to \mathbf{Prs}$ by $\mathcal{2}w = \mathcal{P}ow \circ \mathcal{O}$, etc. Since $\mathcal{O}$ preserves pullback diagrams, it follows from Lemmas 0 and 4.4 that these functors $\mathcal{2}w$, etc., are interpolating. It may be left to the reader to verify that the associated relational extensions are determined in

Lemma 4. *For a relation $R \subseteq A \times B$, we have*

(a) $u \langle \mathcal{2}i^+.R \rangle v \quad \equiv \quad (\forall (p,I) \in u :: (\exists (q,J) \in v :: p = q \ \wedge \ I \langle R \rangle J))$,

(b) $u \langle \mathcal{2}w^+.R \rangle v \quad \equiv \quad (\forall (p,I) \in u :: (\exists (q,J) \in v :: p = q \ \wedge \ I \langle R \rangle J))$
$\wedge \ (\forall (q,J) \in v :: (\exists (p,I) \in u :: p = q \ \wedge \ I \langle R \rangle J))$.

Example 5. In Lemma 4(b), the reader may recognize bisimulation of processes. In fact, a process with actions in a set A can be described as a triple (X,s,x) where $s \in X \to \mathcal{P}ow.(A \times X)$ and $x \in X$, see e.g. [5, 12]. So it is a $\mathcal{2}w$-scheme where functor

$\mathcal{O}$ is given with respect to the set of operators $Op = A$, all of them with arity 1. It follows then from Lemma 4(b) that simulation with respect to $\mathcal{2}w^+$ is the same as bisimulation of processes, as introduced by Park in [11].

Example 6. Bisimulation and similarity. Let $D = \mathbb{N} \cup \{\mathtt{d}, \mathtt{e}\}$ where $\mathbb{N}$ is the set of natural numbers, and $\mathtt{d}$ and $\mathtt{e}$ are two other symbols. Let $\varphi \in D \to \mathcal{P}ow.D$ be given by $\varphi.0 = \emptyset$, and $\varphi.(n+1) = \{n\}$, and $\varphi.\mathtt{d} = \{2k | k \in \mathbb{N}\}$ and $\varphi.\mathtt{e} = \{2k+1 | k \in \mathbb{N}\}$. We consider the $\mathcal{P}ow$-schemes $(D, \varphi, \mathtt{d})$ and $(D, \varphi, \mathtt{e})$. Regarded as processes, $\mathtt{d}$ can perform an arbitrary odd number of transitions and $\mathtt{e}$ can perform an arbitrary even number of transitions. Using Lemma 1 and the definition of simulation, one can show that there is no $\mathcal{P}ow^+$-simulation R with $\mathtt{d}\,\langle R\rangle\,\mathtt{e}$. In terms of process algebras, there is no bisimulation.

We can also regard the schemes over the functor $\mathcal{P}oi$. In that case, we can use relation

$$R = \{(n, n+1) | n \in \mathbb{N}\} \cup \{(\mathtt{d}, \mathtt{e}), (\mathtt{e}, \mathtt{d})\},$$

which is a $\mathcal{P}oi^+$-simulation. Therefore, $\mathtt{d}$ and $\mathtt{e}$ simulate each other with respect to $\mathcal{P}oi^+$.

Since we associate the term bisimulation with the requirement of Lemma 1(c), we shall use the term similarity to express that two schemes simulate each other. Here, $(D, \varphi, \mathtt{d})$ and $(D, \varphi, \mathtt{e})$ are similar for $\mathcal{P}oi^+$. One can also prove that they are similar for $\mathcal{P}oc^+$.

5.4. Functors with bounded spread

We define a functor F to have *spread bounded by* the cardinal number γ iff, for every set X and every element $y \in F.X$, there is a subset $U \subseteq X$ with $\#U \leqslant \gamma$ such that y is in the image of $F.U$ in $F.X$. We define F to have bounded spread iff there is a cardinal number γ such that F has spread bounded by γ. The concept of bounded spread will be useful in our main result, Theorem 6.9.

The functors $\mathcal{O}$, $\mathcal{P}owk$, $\mathcal{P}ownk$, $\mathcal{P}oik$, $\mathcal{P}oink$, $\mathcal{2}ck$, $\mathcal{2}cnk$, $\mathcal{2}wk$, $\mathcal{2}ik$ have bounded spread. In fact, every element $y \in \mathcal{O}.X$ is of the form (p, I) with $p \in Op$ and $I \in \eta.p \to X$. Then $y \in F.U$ for $U = \mathrm{im}.I$. Therefore functor $\mathcal{O}$ has spread bounded by $(\sup\ p \in Op :: \eta.p)$. It is easy to verify that $\mathcal{P}owk$, etc., have spread bounded by κ, and that $\mathcal{2}wk$ and $\mathcal{2}ik$ have spread bounded by $\kappa \times (\sup\ p \in Op :: \eta.p)$.

On the other hand, the functor $\mathcal{P}ow$ is not of bounded spread. In fact, for every cardinal number γ there exists a set X with $\#X > \gamma$. Then $X \in \mathcal{P}ow.X$ is such that $X \in \mathcal{P}ow.U$ implies $\#U > \gamma$ for all U.

6. The construction of universal models

This section is devoted to a construction of universal models. We go back to the setting of Section 3 with a functor $F : \mathbf{Set} \to \mathbf{Set}$ and some relational extension G of

it. The universal model is constructed by means of a premodel. We therefore start to show in Section 6.1 how a premodel is transformed into a model (this is a special case of a standard construction in algebra). The premodel we construct consists of similarity classes of schemes. So, in Section 6.2 we introduce and investigate similarity of schemes. In Section 6.3, saturated coalgebras are introduced and it is shown that every saturated coalgebra gives rise to a flat and separating premodel. Finally, in Section 6.4, we use these results to prove that there is a universal model if F is of bounded spread. The ideas in this section were inspired by the work on final coalgebras in [12].

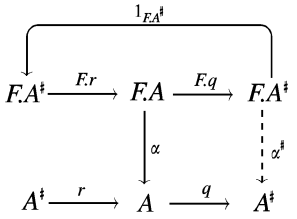
6.1. From premodel to model

Let $(A, \alpha, \leqslant)$ be a G-premodel. We define $\approx$ to be the equivalence relation of A given by

$$x \approx y \quad \equiv \quad x \leqslant y \quad \wedge \quad y \leqslant x.$$

Let $A^\sharp = A/\approx$ be the set of equivalence classes and let $q \in A \to A^\sharp$ be the canonical surjection. We claim that there is a unique function $\alpha^\sharp \in F.A^\sharp \to A^\sharp$ such that $\alpha^\sharp \circ F.q = q \circ \alpha$. To construct $\alpha^\sharp$, we choose a function $r \in A^\sharp \to A$ with $q \circ r = 1_{A^\sharp}$; function r chooses a representing element in each equivalence class. Notice that the functions q and r are both monotonic. We have

$$\begin{aligned}
& q \circ r = 1_{A^\sharp} \\
\Rightarrow\ & \{\text{definitions}\} \\
& (r \circ q, 1_A) \in (=_A) \rightrightarrows (\leqslant) \quad \wedge \quad (1_A, r \circ q) \in (=_A) \rightrightarrows (\leqslant) \\
\Rightarrow\ & \{(\text{rel3}) \text{ and } (\text{rel0})\} \\
& (F.r \circ F.q, 1_{F.A}) \in (=_{F.A}) \rightrightarrows G.(\leqslant) \\
& \quad \wedge (1_{F.A}, F.r \circ F.q) \in (=_{F.A}) \rightrightarrows G.(\leqslant) \\
\Rightarrow\ & \{A \text{ is a } G\text{-premodel}\} \\
& (\alpha \circ F.r \circ F.q, \alpha) \in (=_{F.A}) \rightrightarrows (\leqslant) \cap cv.(\leqslant) \\
\Rightarrow\ & \{\text{definition of } q\} \\
& q \circ \alpha \circ F.r \circ F.q = q \circ \alpha \\
\equiv\ & \{\text{define } \alpha^\sharp = q \circ \alpha \circ F.r\} \\
& \alpha^\sharp \circ F.q = q \circ \alpha.
\end{aligned}$$



So we define $\alpha^\sharp = q \circ \alpha \circ F.r$. The function $\alpha^\sharp$ is uniquely characterized by the last line of this calculation, since $F.q$ is surjective by Lemma 2.0. The preorder $\leqslant$ on A induces an order on the set $A^\sharp$, which is also denoted $\leqslant$. Using Lemma 3.0 on function r, we get that $\alpha^\sharp \in F.A^\sharp \to A^\sharp$ satisfies $(\alpha^\sharp, \alpha^\sharp) \in G.(\leqslant) \rightrightarrows (\leqslant)$. In this way, $(A^\sharp,\ \alpha^\sharp,\ \leqslant)$ is a G-model.

It is easy to verify that, for every coalgebra (D, φ), the functions q and r transfer pre-interpretations and post-interpretations between A and $A^\sharp$. More precisely, if $f \in D \to A$ and $g \in D \to A^\sharp$ are pre-interpretations, then $q \circ f \in D \to A^\sharp$ and $r \circ g \in D \to A$ are pre-interpretations, and similarly for post-interpretations. It follows that if A is flat then $A^\sharp$ is flat, and if A is separating then $A^\sharp$ is separating (see Sections 3.3 and 3.4 for the definitions).

6.2. Simulation and similarity

We come back to relation $\preceq_G$ defined in Section 3.4. We identify a scheme (D, φ, d) with the base point d when the intended coalgebra (D, φ) is clear from the context. In that case, we omit the index G from the symbol $\preceq$. In this way we get, for every pair of coalgebras (D, φ) and (E, ψ), a relation $(\preceq) \subseteq D \times E$ given by

$$x \preceq y \quad \equiv \quad (D, \varphi, x) \preceq_G (E, \psi, y).$$

Lemma 0. *Relation $\preceq$ is the greatest (i.e., weakest) G-simulation between (D, φ) and (E, ψ).*

Proof. By definition, we have $R \subseteq (\preceq)$ for every G-simulation R. Therefore, it suffices to prove that $\preceq$ is a G-simulation, i.e., that $(\varphi, \psi) \in (\preceq) \rightrightarrows G.(\preceq)$.

Let $x \in D,\ y \in E$ with $x \preceq y$. We have to prove $\varphi.x \langle G.(\preceq) \rangle \psi.y$. We can choose a G-simulation R from (D, φ) to (E, ψ) with $x \langle R \rangle y$. Since R is a G-simulation, we have $\varphi.x \langle G.R \rangle \psi.y$. On the other hand, $R \subseteq (\preceq)$ and hence $G.R \subseteq G.(\preceq)$. This implies that $\varphi.x \langle G.(\preceq) \rangle \psi.y$. □

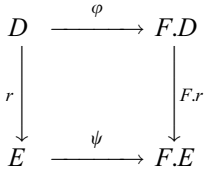
We use relation $\preceq_G$ to define the G-similarity relation $\simeq_G$ between F-schemes by

$$x \simeq_G y \quad \equiv \quad x \preceq_G y \quad \wedge \quad y \preceq_G x.$$

Since $\preceq_G$ is a preorder, relation $\simeq_G$ is an equivalence relation of schemes. When schemes are given by base points, we write $\simeq$ for $\simeq_G$.

Remark. We do not use the term bisimulation here since the relations used to infer $x \preceq_G y$ and $y \preceq_G x$ may be different, whereas bisimulation of processes requires the same relation for both. See Example 6 in Section 5.

Let r be a function from coalgebra (D, φ) to coalgebra (E, ψ). Function r is said to *preserve similarity* iff $x \simeq r.x$ for all $x \in D$. Function r is called a *comorphism* iff $F.r \circ \varphi = \psi \circ r$.



Lemma 1. *Every comorphism preserves similarity.*

Proof. This follows from the next lemma, with the discrete order $=_E$ on E (a direct proof can also be given, but is not much shorter). □

Lemma 2. *Let (E, ψ) be a coalgebra with preorder $\leqslant$ such that $(\psi, \psi) \in (\leqslant) \rightrightarrows G.(\leqslant)$. Let (D, φ) be a coalgebra and let $r \in D \to E$.*
(a) *Assume $F.r \circ \varphi \langle G.(\leqslant) \rangle \psi \circ r$. Then $(D, \varphi, d) \preceq_G (E, \psi, r.d)$ for all $d \in D$.*
(b) *Assume $\psi \circ r \langle G.(\leqslant) \rangle F.r \circ \varphi$. Then $(E, \psi, r.d) \preceq_G (D, \varphi, d)$ for all $d \in D$.*

Proof. (a) Let relation $R \subseteq D \times E$ be defined by $x \langle R \rangle y$ iff $r.x \leqslant y$. We have $d \langle R \rangle r.d$ for all $d \in D$. So, it remains to prove that R is a simulation:

$$
\begin{aligned}
& (\varphi, \psi) \in R \rightrightarrows G.R \\
\equiv\ & \{\text{definitions}\} \\
& (\forall x, y : r.x \leqslant y : \varphi.x \langle G.([1_D, r] \circ (\leqslant)) \rangle \psi.y) \\
\Leftarrow\ & \{\text{calculus and Lemma 4.2}\} \\
& (\forall x, y : r.x \leqslant y : \varphi.x \langle [1_{F.D}, F.r] \circ G.(\leqslant)) \rangle \psi.y) \\
\equiv\ & \{\text{calculus}\} \\
& (\forall x, y : r.x \leqslant y : F.r.(\varphi.x) \langle G.(\leqslant) \rangle \psi.y) \\
\Leftarrow\ & \{\text{transitivity of } G.(\leqslant)\} \\
& (\forall x, y : r.x \leqslant y : F.r.(\varphi.x) \langle G.(\leqslant)) \rangle \psi.(r.x) \quad \wedge \quad \psi.(r.x) \langle G.(\leqslant) \rangle \psi.y)
\end{aligned}
$$

$\equiv$ {assumption of ψ yields second conjunct}

$F.r \circ \varphi \langle G.(\leqslant) \rangle \psi \circ r.$

(b) The other case is similar and uses relation $R \subseteq E \times D$ given by $R = (\leqslant) \circ [r, 1_D]$. □

Let (D, φ) be an F-coalgebra. We shall now use that the pair $(F.D, F.\varphi)$ is also an F-coalgebra. Let us call it the F-transform of (D, φ). Function $\varphi \in D \to F.D$ is a comorphism because of $F.\varphi \circ \varphi = F.\varphi \circ \varphi$. By Lemma 1, this implies that φ preserves similarity.

Lemma 3. *Consider coalgebras (D, φ) and (E, ψ), and their F-transforms.*

(a) $\varphi.d \preceq \psi.e \quad \equiv \quad d \preceq e$ *for all* $d \in D, e \in E$.

(b) $x \langle G.(\preceq) \rangle y \quad \equiv \quad x \preceq y$ *for all* $x \in F.D, y \in F.E$.

Proof. (a) Since $\preceq$ is given by restricting $\preceq_G$, and $\preceq_G$ is a preorder on schemes, this follows from the fact that φ and ψ preserve similarity by Lemma 1. In fact, $\varphi.d \simeq d$ and $\psi.e \simeq e$.

(b) In order to prove the implication ($\Rightarrow$), it suffices to prove that relation $G.(\preceq) \subseteq F.D \times F.E$ is a G-simulation from coalgebra $(F.D, F.\varphi)$ to $(F.E, F.\psi)$, i.e., that

$$(F.\varphi, F.\psi) \in G.(\preceq) \rightrightarrows G.(G.(\preceq)).$$

By (rel3), this follows from Lemma 0.

In order to prove ($\Leftarrow$) of (b), we observe that part (a) implies

$$[1_D, \varphi] \circ (\preceq) \circ [\psi, 1_E] = (\preceq),$$

and hence by (rel2) and Lemma 4.2

$$(*) \qquad [1_{F.D}, F.\varphi] \circ G.(\preceq) \circ [F.\psi, 1_{F.E}] \subseteq G.(\preceq).$$

This enables us to conclude

$x \langle G.(\preceq) \rangle y$

$\Leftarrow$ {$(*)$}

$F.\varphi.x \langle G.(\preceq) \rangle F.\psi.y$

$\Leftarrow$ {Lemma 0 applied to the F-transforms}

$x \preceq y.$ □

6.3. Saturated coalgebras

Lemma 3(a) implies that $\varphi \in D \to F.D$ induces an injection from the set of similarity classes of D to the set of similarity classes of $F.D$. The coalgebra (D, φ) is called saturated if φ induces a bijection. More concretely, the coalgebra (D, φ) is defined to be *G-saturated* iff, for every $x \in F.D$, there exists $y \in D$ with $x \simeq y$.

In the remainder of this subsection, we assume that (H, ρ) is a G-saturated coalgebra. It follows that we can choose a function $\delta \in F.H \to H$ that preserves similarity. In this way, (H, δ) is an F-algebra. It follows from Lemma 1 that

$$\rho.(\delta.x) \simeq x \text{ for all } x \in F.H.$$

For $x, y \in F.H$ we have

$$\begin{aligned}
& \delta.x \preceq \delta.y \\
\equiv\ & \{\text{Lemma 3(a)}\} \\
& \rho.(\delta.x) \preceq \rho.(\delta.y) \\
\equiv\ & \{\text{previous similarity}\} \\
& x \preceq y \\
\equiv\ & \{\text{Lemma 3(b)}\} \\
& x \langle G.(\preceq) \rangle y.
\end{aligned}$$

This proves that the pair $(H, \delta, \preceq)$ is a G-premodel.

Now that we have a G-premodel, we may consider pre-interpretations and post-interpretations in it.

Lemma 4. *Let (D, φ) be a coalgera and let $r \in D \to H$ be a function.*
(a) *If r is a pre-interpretation, then $(D, \varphi, d) \preceq_G (H, \rho, r.d)$ for all $d \in D$.*
(b) *If r is a post-interpretation, then $(H, \rho, r.d) \preceq_G (D, \varphi, d)$ for all $d \in D$.*
(c) *If function r from (D, φ) to (H, ρ) preserves similarity, then $P_\varphi.r \simeq r$.*

Proof.
(a)

$$\begin{aligned}
& P_\varphi.r \preceq r \\
\equiv\ & \{\text{definition of } P_\varphi\} \\
& \delta \circ F.r \circ \varphi \preceq r \\
\equiv\ & \{\text{Lemma 3(a) and } \rho \circ \delta \simeq 1_{F.E}\} \\
& F.r \circ \varphi \preceq \rho \circ r
\end{aligned}$$

$\equiv$ {Lemma 3(b)}

$F.r \circ \varphi \langle G.(\preceq) \rangle \rho \circ r$

$\Rightarrow$ {Lemma 2(a) and Lemma 0}

$(D, \varphi, d) \preceq_G (H, \rho, r.d)$ for all $d \in D$.

(b) This case is similar. It uses Lemma 2(b) instead of 2(a).
(c)

$P_\varphi.r \preceq r$

$\equiv$ {as in the previous cases}

$F.r \circ \varphi \simeq \rho \circ r$

$\Leftarrow$ {$\varphi, \rho,$ and r preserve similarity}

$F.r$ preserves similarity

$\equiv$ {formalization}

$[1_{F.D}, F.r] \subseteq (\preceq) \quad \wedge \quad [F.r, 1_{F.D}] \subseteq (\preceq)$

$\equiv$ {Lemma 3(b)}

$[1_{F.D}, F.r] \subseteq G.(\preceq) \quad \wedge \quad [F.r, 1_{F.D}] \subseteq G.(\preceq)$

$\Leftarrow$ {Lemma 4.2 and (rel1)}

$[1_D, r] \subseteq (\preceq) \quad \wedge \quad [r, 1_D] \subseteq (\preceq)$

$\equiv$ {r preserves similarity}

true. $\square$

Lemma 5. *The G-premodel* $(H, \delta, \preceq)$ *is flat and separating* (*see Sections* 3.3 *and* 3.4).

Proof. First we prove flatness. Let (D, φ) be a coalgebra with a post-interpretation $f \in D \to H$ and a pre-interpretation $g \in D \to H$. Lemma 4 yields

$$(H, \rho, f.d) \preceq_G (D, \varphi, d) \preceq_G (H, \rho, g.d)$$

for all d. By transitivity of $\preceq_G$ and the definition of $\preceq$ on H, this implies $f.d \preceq g.d$ for all $d \in D$, and hence $f \preceq g$. This proves that H is flat.

Secondly, let $f \in D \to H$ be a pre-interpretation of coalgebra (D, φ) and let $g \in E \to H$ be a post-interpretation of coalgebra (E, ψ). Let $d \in D$ and $e \in E$ be such that

$f.d \preceq g.e$. It follows from Lemma 4 that

$$(D,\varphi,d) \preceq_G (H,\rho,f.d) \preceq_G (H,\rho,g.e) \preceq_G (E,\psi,e).$$

This implies $(D,\varphi,d) \preceq_G (E,\psi,e)$, i.e., the existence of a G-simulation of schemes from (D,φ,d) to (E,ψ,e). This proves that H is separating. □

We have thus shown that every saturated coalgebra (H,ρ) gives rise to a flat and separating G-premodel $(H,\delta,\preceq)$. This is the first step in the construction of a universal model, as asked for in Section 3.4.

6.4. The existence of saturated coalgebras

Above we have shown that saturated coalgebras can be useful. This justifies the effort to construct them, if possible. The problem is that, in general, the set $F.D$ is bigger than D, so that the schemes in $F.D$ can be more complex than the schemes in D. The solution lies in the fact that the scheme up to similarity is determined by its shape in the "neighbourhood" of its base point. For this purpose, we introduce the assumption that F is of bounded spread, see Section 5.4.

Lemma 6. *Assume that F has spread bounded by γ. Then every F-scheme (D,φ,d) is similar to an F-scheme (D',φ',d') with $\#D' \leqslant (\gamma+1)^\omega$.*

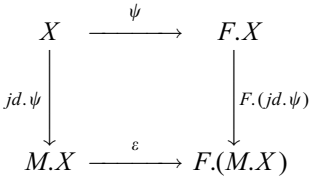
Proof. Let (D,φ,d) be an F-scheme. Since F is of spread bounded by γ, we can choose, for every $x \in D$, a subset $B.x \subseteq D$ such that $\varphi.x$ is in the image of $F.(B.x)$ in $F.D$, and that $x \in B.x$, and that $\#(B.x) \leqslant \gamma+1$.

We now define $B_0 = \{d\}$ and $B_{n+1} = (\bigcup x \in B_n :: B.x)$ for all n, and $D' = (\bigcup n :: B_n)$. For every $x \in D'$, there is n with $x \in B_n$, and hence $B.x \subseteq D'$. It follows that for every $x \in D'$ we have that $\varphi.x$ is in the image of $F.(D')$ in $F.D$. This implies that function $\varphi \in D \to F.D$ restricts to a function $\varphi' \in D' \to F.D'$ such that the inclusion $D' \to D$ is a comorphism from coalgebra (D',φ') to (D,φ). By Lemma 1, therefore, the two schemes (D',φ',d) and (D,φ,d) are similar. Finally, since $\#B.x \leqslant \gamma+1$ for every x, we have $\#B_n \leqslant (\gamma+1)^n$ for all n, and therefore $\#D' \leqslant (\gamma+1)^\omega$. □

The next point is to construct an F-coalgebra that "contains" all coalgebra structures on a given set. This is done as follows.

For a set X, let $M.X$ be the set of pairs (ψ,x) with $\psi \in X \to F.X$ and $x \in X$. For $\psi \in X \to F.X$, we define $jd.\psi \in X \to M.X$ by $jd.\psi.x = (\psi,x)$. We define function $\varepsilon \in M.X \to F.(M.X)$ by $\varepsilon.(\psi,x) = F.(jd.\psi).(\psi.x)$. In this way, obviously, $(M.X,\varepsilon)$ is a coalgebra. The main point of the construction is that, for any $\psi \in X \to F.X$, the function $jd.\psi \in X \to M.X$ is a comorphism from (X,ψ) to $(M.X,\varepsilon)$,

since $\varepsilon \circ jd.\psi = F.(jd.\psi) \circ \psi$.



Lemma 7. *Let (D, φ, d) be an F-scheme and let function $i \in D \to X$ be injective. Then there is $y \in M.X$ such that $(D, \varphi, d) \simeq_G (M.X, \varepsilon, y)$.*

Proof. Let $D' \subseteq X$ be the image of i. Since $i \in D \to D'$ is bijective there is a unique function $\psi' \in D' \to F.X$ with $\psi' \circ i = F.i \circ \varphi$. We define $\psi \in X \to F.X$ by $\psi | D' = \psi'$ and $\psi.x = F.i.(\varphi.d)$ for $x \notin D'$. Then $\psi \circ i = F.i \circ \varphi$. So i is a comorphism from (D, φ) to (X, ψ). The composition $jd.\psi \circ i$ is a comorphism from (D, φ) to $(M.X, \varepsilon)$. Choosing $y = jd.\psi.(i.d)$ we have $(D, \varphi, d) \simeq_G (M.X, \varepsilon, y)$ by Lemma 1. □

Lemma 8. *Assume that the functor F has spread bounded by γ. Let X be a set with cardinality $\geqslant(\gamma + 1)^\omega$.*
(a) *For every F-scheme (D, φ, d) there is $y \in M.X$ such that $(D, \varphi, d) \simeq_G (M.X, \varepsilon, y)$.*
(b) *The coalgebra $(M.X, \varepsilon)$ is G-saturated.*

Proof. (a) This follows from the Lemmas 6 and 7, since for every set D' with $\#D' \leqslant (\gamma + 1)^\omega$ there exists an injective function $i \in D' \to X$.
(b) This follows from part (a) and the definition of saturation. □

As announced in Section 3.4, the existence of a universal model implies that simulation is a complete proof method for ordering in every model. Under a mild condition on the functor involved, we can now formulate and prove our main result that universal models exist.

Theorem 9. *Let the functor F be of bounded spread. Then there is a universal G-model.*

Proof. In view of the final remark in Section 3.4, it suffices to prove the existence of a flat and separating G-model in which every coalgebra has meaning.

Assume F has spread bounded by γ. Choose a set X with cardinality $\geqslant(\gamma+1)^\omega$. By Lemma 8 and the theory of Section 6.3, we can choose $\delta \in F.(M.X) \to M.X$ such that $\varepsilon \circ \delta \simeq 1_{M.X}$ and then $(M.X, \delta, \preceq)$ is a flat and separating G-premodel. Using Section 6.1, we form the G-model $N.X = (M.X^\sharp, \delta^\sharp, \preceq)$, which is also flat and separating.

Let (D, φ) be a coalgebra. By Lemma 8(a), there exists a similarity preserving function $r \in (D, \varphi) \to (M.X, \varepsilon)$. It follows from Lemma 4(c) that the induced function

$r \in D \to N.X$ is an interpretation. Since $N.X$ is flat, this interpretation is the meaning $\mu\varphi$ of (D, φ) in $N.X$. □

Remark. Our universal models play another role than the final coalgebras of [12, 7]. They do not aspire to be the last word on semantics. They only serve to show that if one discards all model assumptions, the order between recursively defined values is due to simulation.

In general, a universal model need not be the algebra obtained from a final coalgebra by reversing the arrow. In fact, for the latter algebra (if it exists), the function $\alpha \in F.A \to A$ is bijective. Therefore, it suffices to give an example where the function α of a universal G-model $(A, \alpha, \leqslant)$ cannot be bijective.

For this purpose, we consider the case defined by the functor $\mathscr{P}oik$ of Section 5.1. So $F = \mathscr{P}oik$ and $G = \mathscr{P}oik^+$. We consider the simple F-schemes $(D, \varphi, \mathtt{d})$ and $(E, \psi, \mathtt{e})$ given by $D = \{\mathtt{d}\}$ and $E = \{\mathtt{e}\}$, and $\varphi.\mathtt{d} = \emptyset$ and $\psi.\mathtt{e} = \{\mathtt{e}\}$. Using Lemma 5.1, we get $(D, \varphi, \mathtt{d}) \preceq_G (E, \psi, \mathtt{e})$ and $(E, \psi, \mathtt{e}) \not\preceq_G (D, \varphi, \mathtt{d})$.

Since $\mathscr{P}oik$ has bounded spread, a universal G-model $(A, \alpha, \leqslant)$ exists (say some $N.X$). Since the model is universal, the two schemes $\mathtt{d}$ and $\mathtt{e}$ mentioned above have meanings $d0 = [\![\mathtt{d}]\!]$ and $e0 = [\![\mathtt{e}]\!]$ in A that satisfy $d0 \leqslant e0$ and $d0 \neq e0$. The sets $\{d0, e0\}$ and $\{e0\}$ are elements of $F.A$ and, by Lemma 5.1, they satisfy $\{d0, e0\} \langle G.(\leqslant) \rangle \{e0\}$, and $\{e0\} \langle G.(\leqslant) \rangle \{d0, e0\}$. Since α is monotonic and $\leqslant$ is an order on A, it follows that $\alpha.\{d0, e0\} = \alpha.\{e0\}$. This proves that α is not injective and hence not bijective.

In passing, one may note that the preorder $G.(\leqslant)$ on $F.A$ is not an order. Also, notice that, if one replaces $\mathscr{P}oik$ by $\mathscr{P}owk$, the example breaks down since then $\mathtt{d} \preceq_G \mathtt{e}$ is invalidated. This implies that the universal model depends on G. □

7. Conclusion and outlook

We have shown that least fixpoint semantics and simulation are closely related, in a very abstract setting. In fact, simulation between recursive definition schemes implies an order relation in almost all models. On the other hand, under mild conditions on the functor involved, there is a model such that, if the values defined are ordered in that model, the schemes have a simulation relation.

The cardinality estimates needed in Section 6 seem to be not essential for the ideas of this paper. It is likely that they can be avoided by leaving ZF set theory and working with classes, as in [1]. Even if that is true, however, it is useful to know that for many functors it is not necessary to leave set theory.

Several further questions emerge. What relations are there with final coalgebras and initial algebras? What happens if we restrict our attention to models that satisfy certain laws? Can we give a similar condition for the order between expressions that contain recursively defined values? What is the impact of the general theory for specific functors and models?

Acknowledgements

We are grateful to Rutger Dijkstra and two referees for questions, suggestions, and criticisms that have led to many improvements.

References

[1] P. Aczel, N. Mendler, A final coalgebra theorem, in: D.H. Pitt, D.E. Rydeheard, P. Dybjer, A.M. Pitts. A. Poigné (Eds.), Category Theory and Computer Science, Proceedings, Manchester, Lecture Notes in Computer Science, 389, Springer, Berlin, 1989, pp. 357–365.
[2] R. Backhouse, J. van der Woude, Lecture Notes of the STOP 1992 Summerschool on Constructive Algorithms, 1992.
[3] G. Birkhoff, On the structure of abstract algebras, Proc. Cambridge Philos. Soc. 31 (1935).
[4] E.W. Dijkstra, C.S. Scholten, Predicate Calculus and Program Semantics, Springer, Berlin, 1990.
[5] W.H. Hesselink, Deadlock and fairness in morphisms of transition systems, Theoret. Comput. Sci. 59 (1988) 235–257.
[6] W.H. Hesselink, Programs, Recursion and Unbounded Choice, Predicate Transformation Semantics and Transformation Rules. Cambridge Tracts in Theoretical Computer Science, vol. 27, Cambridge University Press, Cambridge, 1992.
[7] B. Jacobs, J. Rutten, A tutorial on (co)algebras and (co)induction, Bull. EATCS 62 (1997) 222–259.
[8] E. Meier, Calculating compilers, Thesis Nijmegen, 1992.
[9] R. Milner, An algebraic definition of simulation between programs, Proc. 2nd Internat. Joint Conf. on Artificial Intelligence, British Comp. Soc., 1971.
[10] Y.N. Moschovakis, The logic of functional recursion, in: M.L. Dalla Chiara et al. (Eds.), Logic and Scientific Methods, Kluwer, Dordrcht, 1997, pp. 179–207.
[11] D.M.R. Park, Concurrency and automata on infinite sequences, in: P. Deusen (Ed.), Proc. 5th Gl Conf., Lecture Notes in Computer Science, vol 104, Springer, Berlin, 1981, pp. 167–183.
[12] J.J.M.M. Rutten, D. Turi, On the foundations of final semantics: non-standard sets, metric spaces, partial orders, in: J.W. de Bakker, W.P. de Roever, G. Rozenberg (Eds.), Proc. REX Workshop on Semantics: Foundations and Applications, Lecture Notes in Computer Science, 666, Springer, Berlin, 1993, pp. 477–530.
[13] G. Takeuti, W.M. Zaring, Introduction to Axiomatic Set Theory, Springer, Berlin, 1971.
[14] A.M. Thijs, Simulation and fixpoint semantics, Thesis, Groningen, 1996.